

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:33:37 ; Search time 9.6 Seconds
(without alignments)
30.649 Million cell updates/sec

Title: US-09-909-348-1
Perfect score: 59
Sequence: 1 CEGDSGGPFV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	12	6	Patent No. 5352664-3
2	59	100.0	23	4	US-08-981-088-2
3	59	100.0	23	4	US-08-330-594-8
4	59	100.0	259	4	US-08-944-483-52
5	59	100.0	295	3	US-08-338-368-2
6	59	100.0	376	2	US-08-558-269-10
7	59	100.0	376	4	US-09-410-882-10
8	59	100.0	579	1	US-08-295-411-4
9	59	100.0	579	3	US-08-117-708-14
10	59	100.0	579	5	PCT-US92-10242-4
11	59	100.0	579	5	US-08-955-471-4
12	59	100.0	615	1	US-07-998-972A-3
13	59	100.0	615	1	US-08-463-953-3
14	59	100.0	615	5	PCT-US92-11357-3
15	59	100.0	615	5	US-08-462-261-3
16	59	100.0	622	3	US-08-952-967-8
17	56	94.9	164	4	US-09-518-046-25
18	56	94.9	255	4	US-08-944-483-67
19	56	94.9	256	2	US-09-027-337-3
20	56	94.9	256	4	US-09-644-600-3
21	56	94.9	416	2	US-09-000-846-2
22	55	93.2	15	4	US-08-981-088-3
23	55	93.2	42	1	US-08-293-778-11
24	53	89.8	250	4	US-08-944-483-51
25	53	89.8	261	6	5270178-5
26	53	89.8	261	6	5270178-19
27	53	89.8	261	6	5270178-20

28	53	89.8	261	6	5270178-21	Patent No. 5270178
29	53	89.8	262	1	US-07-720-189-1	Sequence 1, Appli
30	53	89.8	409	4	US-09-065-872-2	Sequence 2, Appli
31	53	89.8	409	4	US-09-667-570A-2	Sequence 2, Appli
32	53	89.8	410	4	US-09-065-872-1	Sequence 1, Appli
33	53	89.8	410	4	US-09-667-570A-1	Sequence 1, Appli
34	53	89.8	419	1	US-08-295-411-1	Sequence 1, Appli
35	53	89.8	419	2	US-08-955-471-1	Sequence 3, Appli
36	53	89.8	419	4	US-09-667-570A-3	Sequence 1, Appli
37	53	89.8	419	5	PCT-US92-10242-1	Sequence 2, Appli
38	53	89.8	460	2	US-08-756-506-4	Sequence 4, Appli
39	53	89.8	460	2	US-08-756-506-4	Sequence 4, Appli
40	53	89.8	460	6	5270178-13	Patent No. 5270178
41	53	89.8	460	6	5270178-14	Patent No. 5270178
42	53	89.8	460	6	5270178-15	Patent No. 5270178
43	53	89.8	460	6	5270178-16	Patent No. 5270178
44	53	89.8	461	6	5225537-2	Patent No. 5225537
45	53	89.8	461	6	5270178-2	Patent No. 5270178

ALIGNMENTS

RESULT 1
5352664-3
; Patent No. 5352664
; APPLICANT: Carney, Darrell H.; Glenn, Kevin C.
; TITLE OF INVENTION: THROMBIN DERIVED POLYPEPTIDES
; COMPOSITIONS AND METHODS FOR USE
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/925,201
; FILING DATE: 31-OCT-1986
; SEQ ID NO:3:
; LENGTH: 12
5352664-3

Query Match 100.0%; Score 59; DB 6; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEGDSGGPFV 10
| | | | | | | | | |
Db 3 CEGDSGGPFV 12

RESULT 2
US-08-981-088-2
; Sequence 2, Application US/08981088
; Patent No. 6146824
; GENERAL INFORMATION:
; APPLICANT: BAR-SHAVIT, RACHEL
; TITLE OF INVENTION: ANTI-METASTATIC AND ANTI-ANGIOGENIC
; PHARMACEUTICAL COMPOSITIONS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,088
; FILING DATE: 27-JAN-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A

REGISTRATION NUMBER: 27,075
REFERENCE/DOCKET NUMBER: JAO 40455
TELEPHONE: (703)-836-6400
TELEFAX: (703)-836-2787
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-981-088-2

Query Match 100.0%; Score 59; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
|||||

Db 14 CEGDSGGPFV 23

RESULT 3
US-08-330-594-8
Sequence 8, Application US/08330594C
Patent No. 6184342

GENERAL INFORMATION:
APPLICANT: CAREY, DARRELL H.
APPLICANT: RAMAKRISHNAN, SHYAM
TITLE OF INVENTION: SYNTHETIC PEPTIDE NEUTROPHIL CELL CHEMOTACTIC AGENTS
FILE REFERENCE: CHBP:002
CURRENT APPLICATION NUMBER: US/08/330,594C
CURRENT FILING DATE: 1994-10-28
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 8
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-08-330-594-8

Query Match 100.0%; Score 59; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
|||||

Db 14 CEGDSGGPFV 23

RESULT 4
US-08-944-483-52
Sequence 52, Application US/08944483
Patent No. 6232456

GENERAL INFORMATION:
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: OF METHODS USEFUL FOR DETECTING AND TREATING DISEASES
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:

INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6232456e
US-08-944-483-52

Query Match 100.0%; Score 59; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
|||||

Db 201 CEGDSGGPFV 210

RESULT 5
US-08-338-368-2
Sequence 2, Application US/08338368
Patent No. 6110721
GENERAL INFORMATION:
APPLICANT: GIBBS, CRAIG S.
APPLICANT: LEUNG, LAWRENCE L.K.
APPLICANT: TSIANG, MANUEL
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND COAGULATION
TITLE OF INVENTION: THERAPY
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:

ADDRESSEE: GILEAD SCIENCES, INC.
STREET: 353 LAKESIDE DRIVE
CITY: FOSTER CITY
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,368
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,038

```

; FILING DATE: 10-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: HENSLEY, MAX D. 27,043
; REGISTRATION NUMBER: 190.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-574-3000
; TELEFAX: 415-573-4899
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-338-368-2

Query Match 100.0%; Score 59; DB 3; Length 295;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 237 CEGDSGGPFV 246

RESULT 6
US-08-558-269-10
; Sequence 10, Application US/08558269
; Patent No. 5961973
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,269
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/847,800
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CRI-001CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-410-882-10

Query Match 100.0%; Score 59; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 318 CEGDSGGPFV 327

RESULT 7
US-09-410-882-10
; Sequence 10, Application US/09410882
; Patent No. 6287561
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/410,882
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,269
; FILING DATE: 13-NOV-1995
; APPLICATION NUMBER: US 07/847,800
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CRI-001CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-410-882-10

Query Match 100.0%; Score 59; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 318 CEGDSGGPFV 327

RESULT 8
US-08-295-411-4
; Sequence 4, Application US/08295411
; Patent No. 5679639
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Mesters, Rolf M.
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; ADDRESSEE: Research Institute
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 No. 5679639th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA

```

COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/295,411
FILING DATE: 22-AUG-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,989
FILING DATE: 18-NOV-1991
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSRI263.0C1
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Region
LOCATION: 1..320
OTHER INFORMATION: /note= "Prothrombin Light Chain"
FEATURE:
NAME/KEY: Region
LOCATION: 321..579
OTHER INFORMATION: /note= "Prothrombin Heavy Chain"
US-08-295-411-4
Query Match 100.0%; Score 59; DB 1; Length 579;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEGDGGPFV 10
Db 521 CEGDGGPFV 530
RESULT 9
US-08-955-471-4
Sequence 4, Application US/08955471
Patent No. 5968751
GENERAL INFORMATION:
APPLICANT: Griffin, John H.
APPLICANT: Mesters, Rolf M.
TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
TITLE OF INVENTION: for Inhibiting Coagulation
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Office of Patent Counsel, The Scripps
ADDRESSEE: Research Institute
STREET: 10666 No. 5968751th Torrey Pines Road, NPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,471
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/295,411
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSRI263.0C1
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Region
LOCATION: 1..320
OTHER INFORMATION: /note= "Prothrombin Light Chain"
FEATURE:
NAME/KEY: Region
LOCATION: 321..579
OTHER INFORMATION: /note= "Prothrombin Heavy Chain"
US-08-955-471-4
Query Match 100.0%; Score 59; DB 2; Length 579;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEGDGGPFV 10
Db 521 CEGDGGPFV 530
RESULT 10
US-09-117-708-14
Sequence 14, Application US/09117708
Patent No. 6060300
GENERAL INFORMATION:
APPLICANT: RADTSHC, Martin; FRIEDRICH, Thomas;
APPLICANT: BOLLSCHEWELER, Claus; SCHMIDT, Martin;
APPLICANT: HOFFEN, Hans Wolfgang; SCHWEDEN, Juergen;
APPLICANT: and RUEBSAMEN, Klaus
TITLE OF INVENTION: Thrombin muteins as antidote for
TITLE OF INVENTION: thrombin inhibitors
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kell & Weinkauff
STREET: 1101 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM AT-compatible, Pentium processor
OPERATING SYSTEM: Windows 95
SOFTWARE: Wordperfect version 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/117,708
FILING DATE: 04-AUG-1998
CLASSIFICATION: 424
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 amino acids

Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
Db 521 CEGDSGGPFV 530

RESULT 12
US-07-998-972A-3
; Sequence 3, Application US/07998972A
; Patent No. 5476777
; GENERAL INFORMATION:
; APPLICANT: Holly, Richard D.
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: METHODS FOR PRODUCING THROMBIN
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Street Tower,
; CITY: Twentieth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/998,972A
; FILING DATE: 19921230
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/860,701
; FILING DATE: 31-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/816,281
; FILING DATE: 31-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-12-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-998-972A-3

Query Match 100.0%; Score 59; DB 1; Length 615;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
Db 557 CEGDSGGPFV 566

RESULT 13
US-08-463-953-3
; Sequence 3, Application US/08463953
; Patent No. 5502034
; GENERAL INFORMATION:
; APPLICANT: Holly, Richard D.
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: METHODS FOR PRODUCING THROMBIN
; NUMBER OF SEQUENCES: 48

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-09-117-708-14

Query Match 100.0%; Score 59; DB 3; Length 579;
Best Local Similarity 100.0%; Pred. No. 0.081;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
Db 521 CEGDSGGPFV 530

RESULT 11
PCT-US92-10242-4.
; Sequence 4, Application PC/TUS9210242
; GENERAL INFORMATION:
; APPLICANT: Griffon, John H.
; APPLICANT: Mesters, Rolf
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10242
; FILING DATE: 19921118
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,989
; FILING DATE: 18-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRO472P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-6312
; TELEFAX: 619-554-2937
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..320
; OTHER INFORMATION: /note= "Prothrombin Light Chain"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 321..579
; OTHER INFORMATION: /note= "Prothrombin Heavy Chain"
PCT-US92-10242-4

Query Match 100.0%; Score 59; DB 5; Length 579;

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend
;; STREET: One Market Plaza, Stewart Street Tower,
;; STREET: Twentieth Floor
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94105
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/463,953
;; FILING DATE:
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/860,701
;; FILING DATE: 31-MAR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/816,281
;; FILING DATE: 31-DEC-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parmelee, Steven W
;; REGISTRATION NUMBER: 31,990
;; REFERENCE/DOCKET NUMBER: 13952-12-2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-467-9600
;; TELEFAX: 415-543-5043
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 615 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-463-953-3

Query Match 100.0%; Score 59; DB 1; Length 615;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDGGGPFV 10
|||||
Db 557 CEGDGGGPFV 566

RESULT 14
US-08-462-261-3
;; Sequence 3, Application US/08462261
;; Patent No. 5527692
;; GENERAL INFORMATION:
;; APPLICANT: Holly, Richard D.
;; APPLICANT: Foster, Donald C.
;; TITLE OF INVENTION: METHODS FOR PRODUCING THROMBIN
;; NUMBER OF SEQUENCES: 48
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend
;; STREET: One Market Plaza, Stewart Street Tower,
;; STREET: Twentieth Floor
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94105
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/462,261
;; FILING DATE: 05-JUN-1995
;; CLASSIFICATION: 424

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/998,972
;; FILING DATE: 30-DEC-1992
;; APPLICATION NUMBER: US 07/860,701
;; FILING DATE: 31-MAR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/816,281
;; FILING DATE: 31-DEC-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parmelee, Steven W
;; REGISTRATION NUMBER: 31,990
;; REFERENCE/DOCKET NUMBER: 13952-12-2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-467-9600
;; TELEFAX: 415-543-5043
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 615 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-462-261-3

Query Match 100.0%; Score 59; DB 1; Length 615;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDGGGPFV 10
|||||
Db 557 CEGDGGGPFV 566

RESULT 15
PCT-US92-11357-3
;; Sequence 3, Application PC/TUS9211357
;; GENERAL INFORMATION:
;; APPLICANT: Holly, Richard D.
;; APPLICANT: Foster, Donald C.
;; TITLE OF INVENTION: METHODS FOR PRODUCING THROMBIN
;; NUMBER OF SEQUENCES: 48
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend
;; STREET: One Market Plaza, Stewart Street Tower,
;; STREET: Twentieth Floor
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94105
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US92/11357
;; FILING DATE: 19921230
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/860,701
;; FILING DATE: 31-MAR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/816,281
;; FILING DATE: 31-DEC-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parmelee, Steven W
;; REGISTRATION NUMBER: 31,990
;; REFERENCE/DOCKET NUMBER: 13952-12-2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-467-9600
;; TELEFAX: 415-543-5043
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 615 amino acids

Tue May 13 09:54:54 2003

; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-11357-3

Query Match 100.0%; Score 59; DB 5; Length 615;
Best Local Similarity 100.0%; Pred. NO. 0.086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
Db 557 CEGDSGGPFV 566

Search completed: May 12, 2003, 15:40:35
Job time : 10.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:33:02 ; Search time 11.4 Seconds
(without alignments)
84.328 Million cell updates/sec

Title: US-09-909-348-1
Perfect score: 59
Sequence: 1 CEGDSGGPFV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	59	100.0	234	2 F42696	thrombin (EC 3.4.21.5)
2	59	100.0	235	2 D42696	thrombin (EC 3.4.21.5)
3	59	100.0	235	2 E42696	thrombin (EC 3.4.21.5)
4	59	100.0	235	2 H42696	thrombin (EC 3.4.21.5)
5	59	100.0	236	2 C42696	thrombin (EC 3.4.21.5)
6	59	100.0	236	2 I42696	thrombin (EC 3.4.21.5)
7	59	100.0	239	2 G42696	thrombin (EC 3.4.21.5)
8	59	100.0	617	2 S10511	thrombin (EC 3.4.21.5)
9	59	100.0	618	2 A35827	thrombin (EC 3.4.21.5)
10	59	100.0	622	1 TBHU	thrombin (EC 3.4.21.5)
11	59	100.0	625	1 TBBO	hepsin (EC 3.4.21.5)
12	56	94.9	417	1 S00845	polyprotein - Afri
13	56	94.9	1524	2 T30337	probable serine pr
14	53	89.8	225	2 S45356	protein C (activat
15	53	89.8	456	1 KXBO	protein C (activat
16	53	89.8	461	1 KXHU	chymotrypsin (EC 3
17	52	88.1	248	2 S49323	hypothetical prote
18	52	88.1	265	2 T15451	factor IX - rabbit
19	52	88.1	275	2 I46712	coagulation factor
20	52	88.1	282	2 I84621	coagulation factor
21	52	88.1	459	2 J00419	hypodermin C (EC 3
22	51	86.4	230	2 A27802	serine proteinase
23	50	84.7	121	2 PS0049	chymotrypsin-like
24	50	84.7	126	2 A23473	chymotrypsin (EC 3
25	50	84.7	218	1 KYVH2C	trypsin (EC 3.4.21
26	50	84.7	237	2 S68702	trypsin-like prote
27	50	84.7	238	1 TRWVS1	trypsin-like prote
28	50	84.7	254	1 TRWVS1	trypsin-like prote
29	50	84.7	254	2 S65465	trypsin-like prote

30	50	84.7	256	1 TRFF	trypsin-like prote
31	50	84.7	260	2 I56559	neuropsin - mouse
32	50	84.7	264	2 I38136	chymotrypsin-like
33	50	84.7	264	2 S32794	trypsin-like prote
34	50	84.7	267	2 S40006	trypsin (EC 3.4.21
35	50	84.7	270	2 S56160	mast cell tryptase
36	50	84.7	271	2 S41308	serine proteinase
37	50	84.7	272	2 JC4170	trypsin-like prote
38	50	84.7	273	2 A47246	trypsin (EC 3.4.21
39	50	84.7	274	2 S35339	trypsin (EC 3.4.21
40	50	84.7	274	2 JC4171	trypsin (EC 3.4.21
41	50	84.7	275	2 S40007	trypsin (EC 3.4.21
42	50	84.7	275	2 S40005	trypsin (EC 3.4.21
43	50	84.7	275	2 A35863	trypsin (EC 3.4.21
44	50	84.7	275	2 B35863	trypsin (EC 3.4.21
45	50	84.7	275	2 C35863	trypsin (EC 3.4.21

ALIGNMENTS

RESULT 1
F42696
thrombin (EC 3.4.21.5) B chain - Cynops pyrogastror (fire-bellied newt) (fragment)
C:Species: Cynops pyrogastror (fire-bellied newt)
C>Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 17-Mar-1999
C:Accession: F42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Note: sequence not
A:Accession: F42696
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-234 <BAN>
A:Cross-references: GB:M81395
A:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase

Query Match 100.0%; Score 59; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 0.0092;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
| | | | | | | | | |
Db 176 CEGDSGGPFV 185

RESULT 2
D42696
thrombin (EC 3.4.21.5) B chain - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
C:Accession: D42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-235 <BAN>
A:Cross-references: GB:M81391
A:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 59; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10

```
|||||
Db 177 CEGDSGGPFV 186

RESULT 3
E42696
thrombin (EC 3.4.21.5) B chain - tokay (fragment)
C:Species: gekko gekko (tokay)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
C:Accession: E42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: E42696
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-235 <BAN>
A:Cross-references: GB:M81392
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 59; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
|||||
Db 177 CEGDSGGPFV 186

RESULT 4
H42696
thrombin (EC 3.4.21.5) B chain - white sturgeon (fragment)
C:Species: Acipenser transmontanus (white sturgeon)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
C:Accession: H42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: H42696
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-235 <BAN>
A:Cross-references: GB:M81399
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 59; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
|||||
Db 177 CEGDSGGPFV 186

RESULT 5
C42696
thrombin (EC 3.4.21.5) B chain - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
C:Accession: C42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: C42696
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA

A:Residues: 1-236 <BAN>
A:Cross-references: GB:M81396
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-227/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 59; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
|||||
Db 178 CEGDSGGPFV 187

RESULT 6
I42696
thrombin (EC 3.4.21.5) B chain - Pacific hagfish (fragment)
C:Species: Eptatretus stouti (Pacific hagfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
C:Accession: I42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: I42696
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-236 <BAN>
A:Cross-references: GB:M81393
A:Note: nucleotide translation not given
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 59; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
|||||
Db 177 CEGDSGGPFV 186

RESULT 7
G42696
thrombin (EC 3.4.21.5) B chain - rainbow trout (fragment)
C:Species: Oncorhynchus mykiss (rainbow trout)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: G42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequence
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: G42696
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-239 <BAN>
A:Cross-references: GB:M81398; NID:g213486; PIDN:AAA49433.1; PID:g213487
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 100.0%; Score 59; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 0.0094;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
|||||
Db 177 CEGDSGGPFV 186

RESULT 8
```

Tue May 13 09:54:58 2003

S10511
thrombin (EC 3.4.21.5) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 03-May-2002
C:Accession: S10511; A60576; B42696
R:Phanich, M.; Monard, D.
Nucleic Acids Res. 18, 4251, 1990
A:Title: cDNA sequence of rat prothrombin.
A:Reference number: S10511; MUID:90332426; PMID:2377469
A:Accession: S10511
A:Molecule type: mRNA
A:Residues: 1-617 <DIH>
A:Cross-references: EMBL:X52835; NID:956969; PIDN:CAA37017.1; PID:956970
R:Henrikson, K.P.; Jazin, E.E.; Greenwood, J.A.; Dickerman, H.W.
Endocrinology 126, 167-175, 1990
A:Title: Prothrombin levels are increased in the estrogen-treated immature rat uterus.
A:Reference number: A60576; MUID:90091942; PMID:2293980
A:Accession: A60576
A:Molecule type: protein
A:Residues: 44-58 <HEN>
A:Note: the authors purified the proenzyme from the estrogen-stimulated maturing rat uterus.
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequencing.
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: B42696
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 383-617, 'E' <BAN>
A:Cross-references: GB:M81397
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; hydrolysis
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-43/Domain: propeptide #status predicted <PRO>
F:28-88/Domain: Gla domain homology <GLA>
F:44-617/Product: prothrombin #status experimental <PMAT>
F:109-187/Domain: kringle homology <KR1>
F:215-292/Domain: kringle homology <KR2>
F:360-609/Domain: trypsin homology <TRY>
F:501-5158,60,63,64,69,70,73,76/Modified site: gamma-carboxylglutamic acid (Glu) #status
F:61-66,91-104,109-187,130-170,158-182,215-292,236-276,264-287,332-478,387-403,532-546,548,554,564/Active site: His, Asp, Ser #status predicted
F:402,458,564/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 59; DB 2; Length 617;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 560 CEGDSGGPFV 569

RESULT 9
A35827
thrombin (EC 3.4.21.5) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 03-May-2002
C:Accession: A35827; A42696; S12081
R:Degen, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pal, J.A.
DNA Cell Biol. 9, 487-498, 1990
A:Title: Characterization of the cDNA coding for mouse prothrombin and localization of the gene.
A:Reference number: A35827; MUID:91025551; PMID:2222810
A:Accession: A35827
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-618 <DEG>
A:Cross-references: GB:X52308; NID:953813; PIDN:CAA35548.1; PID:953814
A:Experimental source: strain C57BL/6
A:Note: the data were obtained from females resulting from the cross of M. domesticus and M. musculus.
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequencing.
A:Reference number: A42696; MUID:92212913; PMID:1557383

A:Accession: A42696
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 384-618, 'E' <BAN>
A:Cross-references: GB:M81394
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; hydrolysis
F:1-24/Domain: signal sequence #status predicted <PRO>
F:25-43/Domain: propeptide #status predicted <PRO>
F:28-88/Domain: Gla domain homology <GLA>
F:44-618/Product: prothrombin #status predicted <PMAT>
F:109-187/Domain: kringle homology <KR1>
F:215-293/Domain: kringle homology <KR2>
F:361-610/Domain: trypsin homology <TRY>
F:50,51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxylglutamic acid (Glu) #status
F:61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-548,554,565/Active site: His, Asp, Ser #status predicted
F:403,459,565/Active site: His, Asp, Ser #status predicted

Query Match 100.0%; Score 59; DB 2; Length 618;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 561 CEGDSGGPFV 570

RESULT 10
TBH0
thrombin (EC 3.4.21.5) precursor [validated] - human
N:Alternate names: coagulation factor II
N:Contains: prothrombin
C:Species: Homo sapiens (man)
C:Date: 30-Nov-1980 #sequence_revision 22-Jul-1994 #text_change 08-Dec-2000
C:Accession: A29351; A00914; B00914; A37549; A37550; I51952
R:Degen, S.J.F.; Davie, E.W.
Biochemistry 26, 6165-6177, 1987
A:Title: Nucleotide sequence of the gene for human prothrombin.
A:Reference number: A29351; MUID:88077877; PMID:2825773
A:Accession: A29351
A:Molecule type: DNA
A:Residues: 1-622 <DEG>
A:Cross-references: GB:M17262; GB:M33691; NID:9558069; PIDN:AA63054.1; PID:g339641
R:Degen, S.J.F.; MacGillivray, R.T.A.; Davie, E.W.
Biochemistry 22, 2087-2097, 1983
A:Title: Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin.
A:Reference number: A00914; MUID:83231469; PMID:6303407
A:Accession: A00914
A:Molecule type: mRNA
A:Residues: 8-163, 'N' 165-622 <DE2>
A:Cross-references: GB:V00595; GB:J00307; NID:g37128; PIDN:CAA23842.1; PID:g1335344
A:Accession: B00914
A:Molecule type: DNA
A:Residues: 188-311 <DE3>
R:Walz, D.A.; Hewett-Emmett, D.; Seegers, W.H.
Proc. Natl. Acad. Sci. U.S.A. 74, 1969-1972, 1977
A:Reference number: A37549; MUID:77193964; PMID:266717
A:Accession: A37549
A:Molecule type: protein
A:Residues: 44-118, 'N' 120, 'S' 122-163, 'I' 165-175, 'A' 177-182, 'T' 184-193, 'MV' 196-3
R:Butkowsky, R.J.; Ellison, J.; Downing, M.R.; Mann, K.G.
J. Biol. Chem. 252, 4942-4957, 1977
A:Title: Primary structure of human prothrombin 2 and alpha-thrombin.
A:Reference number: A37550; MUID:77207112; PMID:873923
A:Accession: A37550
A:Molecule type: protein
A:Residues: 315-334, 'N' 336-348, 'N' 350-368, 'N' 370-397, 'N' 399-413, 'N' 415-484, 'N' 4
R:Rabiet, M.J.; Blashill, A.; Furie, B.; Furie, B.C.
J. Biol. Chem. 261, 13210-13215, 1986
A:Reference number: A37551; MUID:87008532; PMID:3759958
A:Contents: annotation; activation cleavages
R:MacGillivray, R.T.; Irwin, D.M.; Guinto, E.R.; Stone, J.C.
Ann. N. Y. Acad. Sci. 485, 73-79, 1986

A:Title: Recombinant genetic approaches to functional mapping of thrombin.
 A:Reference number: 151952; MUID:87182874; PMID:3471151
 A:Accession: 151952

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2, 'R', 5-100 <RES>
 A:Cross-references: GB:M33031; NID:q190723; PIDN:AAA60220.1; PID:q190724

C:Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibrin. Thrombin is activated on the surface of a phospholipid membrane that binds to fibrinogen. Thrombin is activated on the surface of a phospholipid membrane that binds to fibrinogen. Thrombin is activated on the surface of a phospholipid membrane that binds to fibrinogen.

C:Comment: The cleavage after Arg-198, observed in vitro, does not occur in plasma.

C:Comment: The gamma-carboxyglutamyl residues bind calcium ions, result from the carboxyl interaction with the negatively charged phospholipid membrane surface.

C:Comment: The prothrombin precursor is synthesized in the liver.

C:Genetics:

A:Gene: GDB:F2

A:Cross-references: GDB:119894; OMIM:176930

A:Map position: 11p11-11q12

A:Introns: 27/1; 80/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2; 433/2; 491/2; 552/2

C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C:Keywords: acute phase; blood coagulation; calcium binding; carboxylglutamic acid; duplication

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-43/Domain: propeptide #status predicted <PRO>

F:28-87/Domain: Gla domain homology <GLA>

F:44-327/Domain: activation peptide #status experimental <APT>

F:108-186/Domain: kringle homology <KR1>

F:213-291/Domain: kringle homology <KR2>

F:328-363/Product: thrombin light chain #status experimental <LCH>

F:364-622/Product: thrombin heavy chain #status experimental <HCH>

F:364-613/Domain: trypsin homology <TRY>

F:49-50-57-59-62-63-68-69-72-75/Modified site: gamma-carboxyglutamic acid (Glu) #status

F:60-65-90-103-108-186-189-197-181-213-291-234-274-262-286/Disulfide bonds: #status

F:121-143/Binding site: carboxylate (Asn) (covalent) #status predicted

F:336-482-536-550-564-594/Disulfide bonds: #status predicted

F:391-407/Disulfide bonds: #status experimental

F:406-462/Active site: His, Asp #status predicted

F:416/Binding site: carboxylate (Asn) (covalent) #status experimental

F:568/Active site: Ser #status experimental

Query Match 100.0%; Score 59; DB 1; Length 622;

Best Local Similarity 100.0%; Pred. No. 0.024;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10

Db 564 CEGDSGGPFV 573

RESULT 11

TB80

thrombin (EC 3.4.21.5) precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 24-Apr-1984 #sequence_revision 14-Jul-1994 #text_change 18-Jun-1999

C:Accession: S02537; A00915; A37552; I46045; S67518

C:Irwin, D.M.; Robertson, K.A.; MacGillivray, R.T.A.

J. Mol. Biol. 200, 31-45, 1988

A:Title: Structure and evolution of the bovine prothrombin gene.

A:Reference number: S02537; MUID:88245190; PMID:3379642

A:Accession: S02537

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-625 <IR>

R:MacGillivray, R.T.A.; Davie, E.W.

Biochemistry 23, 1626-1634, 1984

A:Title: Characterization of bovine prothrombin mRNA and its translation product.

A:Reference number: A00915; MUID:84203525; PMID:6326805

A:Accession: A00915

A:Molecule type: mRNA

A:Residues: 1-230, H, 232-625 <MAC>

A>Note: 600-Asn was also found

R:Magnusson, S.; Sottrup-Jensen, L.; Petersen, T.E.; Claeys, H.

in Boerhaave Symposium on Prothrombin and Related Coagulation Factors, Hemker, H.C.;
 A:Reference number: A37552
 A:Accession: A37552

A:Molecule type: protein

A:Residues: 44-287, 'N', 289-352, 'E', 354, 'O', 356-548, 'ND', 551-599, 'N', 601-625 <MAC>

A>Note: the evidence for 231-Ser is strong

A>Note: disulfide bonds and carbohydrate binding sites were determined

R:Park, C.H.; Tulinsky, A.

Biochemistry 25, 3977-3982, 1986

A:Title: Three-dimensional structure of the kringle sequence: structure of prothrombin

A:Reference number: A37553; MUID:86296631; PMID:3741841

A:Contents: annotation; residues 44-317, X-ray crystallography, 2.8 angstroms

R:Irwin, D.M.; Ahern, K.G.; Pearson, G.D.; MacGillivray, R.T.A.

Biochemistry 24, 6854-6861, 1985

A:Title: Characterization of the bovine prothrombin gene.

A:Reference number: A37554; MUID:86077733; PMID:3000440

A:Contents: annotation; gene structure

R:MacGillivray, R.T.; Degen, S.J.; Chandra, T.; Woo, S.L.; Davie, E.W.

Proc. Natl. Acad. Sci. U.S.A. 77, 5153-5157, 1980

A:Title: Cloning and analysis of a cDNA coding for bovine prothrombin.

A:Reference number: I46045; MUID:81054926; PMID:6254059

A:Accession: I46045

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 466-599, 'N', 601-625 <MA2>

A:Cross-references: EMBL:V00135; NID:9772; PIDN:CAA23451.1; PID:9808945

R:Pejler, G.; Karlstroem, A.R.; Berg, L.

Eur. J. Biochem. 227, 102-107, 1995

A:Title: Identification of the proteolytic thrombin fragments formed after cleavage w

A:Reference number: S67518; MUID:95154277; PMID:7851376

A:Accession: S67518

A:Status: preliminary

A:Molecule type: protein

A:Residues: 318-325; 333-338, 'X', 340; 367-374; 481-484, 'X', 486-488; 515-522 <PEJ>

C:Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Comment: Thrombin is activated on the surface of a phospholipid membrane that bi

C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 18-Jun-1999
C:Accession: S00845
R:Leytus, S.P.; Loeb, K.R.; Hagen, F.S.; Kurachi, K.; Davie, E.W.
Biochemistry 27, 1067-1074, 1988
A:Title: A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain
A:Reference number: S00845; MUID:88209431; PMID:2835076
A:Accession: S00845
A:Molecule type: mRNA
A:Residues: 1-417 <LEY>
A:Cross-references: EMBL:X07732; NID:g32063; PIDN:CAA30558.1; PID:g32064
C:Genetics:
A:Gene: GDB:HPN; TMPSR31; hepsin
A:Cross-references: GDB:I35685; OMIM:142440
A:Map position: 19q11-19q13.2
C:Superfamily: hepsin; trypsin homology
C:Keywords: hydrolase; liver; serine proteinase; transmembrane protein
F:23-45/Domain: transmembrane #status predicted <TN>
F:163-400/Domain: trypsin homology <TRY>
F:188-204,291-359,322-338,349-381/Disulfide bonds: #status predicted
F:203,257,353/Active site: His, Asp, Ser #status predicted
Query Match 94.9%; Score 56; DB 1; Length 417;
Best Local Similarity 90.0%; Pred. No. 0.052;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEGDSGGPFV 10
I:|||||
Db 349 CQDGGGPFV 358
I:|||||
RESULT 13
T30337
Polyprotein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30337
R:Yang, J.C.; Lindsay, L.L.; Hedrick, J.L.
submitted to the EMBL Data Library, March 1998
A:Description: cDNA cloning of ovochymase, a chymotrypsin-like protease released from X
A:Reference number: Z30829
A:Accession: T30337
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1524 <YAN>
A:Cross-references: EMBL:U81290; NID:g2981640; PID:g2981641; PIDN:AAC24717.1
Query Match 94.9%; Score 56; DB 2; Length 1524;
Best Local Similarity 90.0%; Pred. No. 0.18;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEGDSGGPFV 10
I:|||||
Db 243 CQDGGGPFV 252
I:|||||
RESULT 14
S45356
probable serine proteinase (EC 3.4.21.-) ACO - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 14-Apr-1998 #sequence_revision 08-May-1998 #text_change 19-May-2000
C:Accession: S45356
R:Dihanich, M.; Spiess, M.
Biochim. Biophys. Acta 1218, 225-228, 1994
A:Title: A novel serine proteinase-like sequence from human brain.
A:Reference number: S45356; MUID:94289486; PMID:8018728
A:Accession: S45356
A:Molecule type: mRNA
A:Residues: 1-225 <DIH>
A:Cross-references: EMBL:X75363; NID:g407137; PIDN:CAA53145.1; PID:g940540
A:Experimental source: Alzheimer's disease patient brain cortex
C:Genetics:
A:Gene: ACO
A:Introns: 175/3

C:Superfamily: trypsin; trypsin homology
C:Keywords: Hydrolase; serine proteinase

Query Match 89.8%; Score 53; DB 2; Length 225;
Best Local Similarity 90.0%; Pred. No. 0.095;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
I:|||||
Db 174 CEGDSGGPLV 183
I:|||||

RESULT 15

KXBO

protein C (activated) (EC 3.4.21.69) precursor - bovine (fragment)
N:Alternate names: autoprothrombin IIA; plasma protein C

C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Nov-1980 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999

C:Accession: A26250; A18385; A18386; A00928

R:Long, G.L.; Balagajar, R.M.; MacGillivray, R.T.A.

Proc. Natl. Acad. Sci. U.S.A. 81, 5653-5656, 1984
A:Title: Cloning and sequence of liver cDNA coding for bovine protein C.

A:Reference number: A26250; MUID:85014826; PMID:6091100

A:Accession: A26250

A:Molecule type: mRNA

A:Residues: 1-456 <LON>

R:Fernlund, P.; Stenflo, J.

J. Biol. Chem. 257, 12170-12179, 1982

A:Title: Amino acid sequence of the light chain of bovine protein C.

A:Reference number: A18385; MUID:83007325; PMID:6896876

A:Accession: A18385

A:Molecule type: protein

A:Residues: 40-194 <FER>

A:Note: 82-Lys was also found

R:Drakenberg, T.; Fernlund, P.; Roepstorff, P.; Stenflo, J.

Proc. Natl. Acad. Sci. U.S.A. 80, 1802-1806, 1983

A:Title: beta-Hydroxyaspartic acid in vitamin K-dependent protein C.

A:Reference number: A19316; MUID:83169769; PMID:6572939

A:Contents: annotation; revision to residue 110

R:Stenflo, J.; Fernlund, P.

J. Biol. Chem. 257, 12180-12190, 1982

A:Title: Amino acid sequence of the heavy chain of bovine protein C.

A:Reference number: A18386; MUID:83007326; PMID:6896877

A:Accession: A18386

A:Molecule type: protein

A:Residues: 197-454, 'PV' <STE>

R:Esmon, N.L.; Debault, L.E.; Esmon, C.T.

J. Biol. Chem. 258, 5548-5553, 1983

A:Title: Proteolytic formation and properties of gamma-carboxyglutamic acid-domains

A:Reference number: A37541; MUID:83213513; PMID:6304092

A:Contents: annotation; activation; calcium binding

R:Johnson, A.E.; Esmon, N.L.; Laue, T.N.; Esmon, C.T.

J. Biol. Chem. 258, 5554-5560, 1983

A:Title: Structural changes required for activation of protein C are induced by Ca2+

A:Reference number: A37542; MUID:83213514; PMID:6406503

A:Contents: annotation; activation; calcium binding

C:Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that

s.

C:Comment: Protein C is synthesized in the liver as a single chain precursor, which i

bin, which cleaves a tetradecapeptide from the amino end of the heavy chain; this rea

C:Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with s

cognition of the thrombin-thrombomodulin complex.

C:Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vita

C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol

C:Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium bind

F:1-29/Domain: signal sequence (fragment) #status predicted <SIG>

F:24-83/Domain: Gla domain homology <Gla>

F:30-39/Domain: propeptide #status predicted <PRO>

F:40-194/Product: protein c light chain #status experimental <LCH>

F:98-128/Domain: EGF homology <EG1>

F:137-172/Domain: EGF homology <EG2>

F:197-456/Product: protein C heavy chain #status experimental <HCH>

F:197-210/Domain: activation peptide #status experimental <APT>

F:211-440/Domain: trypsin homology <TRY>
F:45,46,53,55,58,59,62,64,65,68,74/Modified site: gamma-carboxyglutamic acid (Glu) #stat
F:110/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F:119-128,137-148,144-157,159-172,180-318,237-253,368-382,393-421/Disulfide bonds: #stat
F:136,289,350/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:252,298,397/Active site: His, Asp, Ser #status predicted
F:366/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.8%; Score 53; DB 1; Length 456;
Best Local Similarity 90.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
|||||||
Db 393 CEGDSGGPNV 402

Search completed: May 12, 2003, 15:39:40
Job time : 12.4 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:38:52 ; Search time 19.4 Seconds
(without alignments)
47.436 Million cell updates/sec

Title: US-09-909-348-1
Perfect score: 59
Sequence: 1 CEGDSGGPFV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pdb.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pdb.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pdb.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pdb.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pdb.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pdb.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pdb.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pdb.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pdb.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pdb.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pdb.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	10	9	US-10-050-692-1
2	59	100.0	10	9	US-10-050-688-1
3	59	100.0	10	10	US-09-909-348-1
4	59	100.0	10	10	US-09-909-122-1
5	59	100.0	12	9	US-10-050-611-2
6	59	100.0	12	10	US-09-904-090-2
7	59	100.0	23	9	US-10-050-692-6
8	59	100.0	23	9	US-10-050-611-3
9	59	100.0	23	9	US-10-050-611-4
10	59	100.0	23	9	US-10-050-688-5
11	59	100.0	23	9	US-10-050-688-6
12	59	100.0	23	10	US-09-777-328-8
13	59	100.0	23	10	US-09-904-090-3
14	59	100.0	25	9	US-10-050-692-5
15	59	100.0	25	10	US-09-909-348-5
16	59	100.0	25	10	US-09-909-122-5
17	59	100.0	250	9	US-09-898-837A-45
18	59	100.0	251	9	US-09-898-837A-41
19	56	94.9	236	9	US-09-898-837A-44

20	56	94.9	376	10	US-09-820-002-2	Sequence 2, Appli
21	56	94.9	417	9	US-10-073-060-2	Sequence 2, Appli
22	56	94.9	417	10	US-09-820-002-4	Sequence 4, Appli
23	53	89.8	320	10	US-09-888-615-90	Sequence 90, Appl
24	53	89.8	326	10	US-09-888-615-105	Sequence 105, App
25	53	89.8	419	9	US-10-182-263-1	Sequence 1, Appli
26	53	89.8	419	9	US-10-182-263-3	Sequence 3, Appli
27	53	89.8	419	9	US-10-182-263-4	Sequence 4, Appli
28	53	89.8	419	9	US-10-182-263-5	Sequence 5, Appli
29	53	89.8	419	9	US-10-182-263-6	Sequence 6, Appli
30	53	89.8	419	9	US-09-978-917A-4	Sequence 4, Appli
31	53	89.8	461	9	US-10-182-263-2	Sequence 2, Appli
32	53	89.8	461	9	US-09-978-917A-2	Sequence 2, Appli
33	52	88.1	9	9	US-10-102-283-163	Sequence 163, App
34	52	88.1	9	9	US-10-135-795-163	Sequence 163, App
35	52	88.1	9	10	US-09-919-048-163	Sequence 163, App
36	50	84.7	12	9	US-10-125-459-18	Sequence 18, Appl
37	50	84.7	12	9	US-10-125-459-19	Sequence 19, Appl
38	50	84.7	12	9	US-10-067-761-33	Sequence 33, Appl
39	50	84.7	12	9	US-10-067-761-34	Sequence 34, Appl
40	50	84.7	12	10	US-09-879-792-16	Sequence 16, Appl
41	50	84.7	12	10	US-09-804-156-33	Sequence 33, Appl
42	50	84.7	12	10	US-09-804-156-34	Sequence 34, Appl
43	50	84.7	12	10	US-09-946-633-18	Sequence 18, Appl
44	50	84.7	12	10	US-09-946-633-19	Sequence 19, Appl
45	50	84.7	18	9	US-09-826-290-247	Sequence 247, App

ALIGNMENTS

RESULT 1
US-10-050-692-1
; Sequence 1, Application US/10050692
; Publication No. US20020182205A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Grother, Roger S.
; APPLICANT: Simmons, David J.
; APPLICANT: Yang, Jinping
; APPLICANT: Redin, William R.
; TITLE OF INVENTION: STIMULATION OF BONE GROWTH WITH THROMBIN
; TITLE OF INVENTION: PEPTIDE DERIVATIVES
; FILE REFERENCE: 3033, 1002-004
; CURRENT APPLICATION NUMBER: US/10/050,692
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 09/909,122
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/219,300
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human prothrombin
US-10-050-692-1

Query Match 100.0%; Score 59; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEGDSGGPFV 10
| | | | | | | | | |
Db 1 CEGDSGGPFV 10

RESULT 2
US-10-050-688-1
; Sequence 1, Application US/10050688
; Publication No. US20020198154A1

```
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Stiernberg, Janet
; APPLICANT: Bergmann, John
; TITLE OF INVENTION: STIMULATION OF CARTILAGE GROWTH WITH
; TITLE OF INVENTION: AGONISTS OF THE NON-PROTEOLYTICALLY ACTIVATED THROMBIN
; TITLE OF INVENTION: RECEPTOR
; FILE REFERENCE: 3033.1003-004
; CURRENT APPLICATION NUMBER: US/10/050,688
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 09/909,348
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/219,800
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide fragment of thrombin
US-10-050-688-1

Query Match      100.0%; Score 59; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
Db 1 CEGDSGGPFV 10
|||||

RESULT 3
US-09-909-348-1
; Sequence 1, Application US/09909348
; Patent No. US20020042373A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Stiernberg, Janet
; APPLICANT: Bergmann, John
; TITLE OF INVENTION: Stimulation Of Cartilage Growth With Agonists
; TITLE OF INVENTION: Of the No. US20020042373A1-Proteolytically Activated Thrombin Re
; FILE REFERENCE: 3033.1003-001
; CURRENT APPLICATION NUMBER: US/09/909,348
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/219,800
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide fragment of Thrombin
US-09-909-348-1

Query Match      100.0%; Score 59; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
Db 1 CEGDSGGPFV 10
|||||

RESULT 4
US-09-909-122-1
; Sequence 1, Application US/09909122
; Patent No. US20020128202A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Simmons, David J.
; APPLICANT: Yang, Jiping
; APPLICANT: Redin, William R.
; TITLE OF INVENTION: Stimulation Of Bone Growth With Thrombin
; TITLE OF INVENTION: Peptide Derivatives
; FILE REFERENCE: 3033.1002-001
; CURRENT APPLICATION NUMBER: US/09/909,122
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/219,300
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide fragment of Thrombin
US-09-909-122-1

Query Match      100.0%; Score 59; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
Db 1 CEGDSGGPFV 10
|||||

RESULT 5
US-10-050-611-2
; Sequence 2, Application US/10050611
; Publication No. US20020187933A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: 3033.1000-008
; CURRENT APPLICATION NUMBER: US/10/050,611
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 09/904,090
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/217,583
; PRIOR FILING DATE: 2000-07-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human fragment of thrombin
US-10-050-611-2

Query Match      100.0%; Score 59; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
Db 3 CEGDSGGPFV 12
|||||

RESULT 6
US-09-904-090-2
; Sequence 2, Application US/09904090
; Patent No. US20020061852A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell
; TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
```

;; TITLE OF INVENTION: PEPTIDES
;; FILE REFERENCE: 3033.1000-001
;; CURRENT APPLICATION NUMBER: US/09/904,090
;; CURRENT FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: US 60/217,583
;; PRIOR FILING DATE: 2000-07-12
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 12
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Polypeptide, fragment of thrombin
US-09-904-090-2

Query Match 100.0%; Score 59; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
Db 3 CEGDSGGPFV 12

RESULT 7
US-10-050-692-6

;; Sequence 6, Application US/10050692
;; Publication No. US20020182205A1
;; GENERAL INFORMATION:
;; APPLICANT: Carney, Darrell H.
;; APPLICANT: Crowther, Roger S.
;; APPLICANT: Simmons, David J.
;; APPLICANT: Yang, Jinping
;; APPLICANT: Redin, William R.
;; TITLE OF INVENTION: STIMULATION OF BONE GROWTH WITH THROMBIN
;; TITLE OF INVENTION: PEPTIDE DERIVATIVES
;; FILE REFERENCE: 3033.1002-004
;; CURRENT APPLICATION NUMBER: US/10/050,692
;; CURRENT FILING DATE: 2002-01-16
;; PRIOR APPLICATION NUMBER: 09/909,122
;; PRIOR FILING DATE: 2001-07-19
;; PRIOR APPLICATION NUMBER: 60/219,300
;; PRIOR FILING DATE: 2000-07-19
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 6
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: c-terminal amidated fragment of human thrombin
;; FEATURE:
;; NAME/KEY: AMIDATION
;; LOCATION: (23)...(23)
;; OTHER INFORMATION: valine is amidated as CONH2
US-10-050-692-6

Query Match 100.0%; Score 59; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
Db 14 CEGDSGGPFV 23

RESULT 8

US-10-050-611-3
;; Sequence 3, Application US/10050611
;; Publication No. US20020187933A1
;; GENERAL INFORMATION:
;; APPLICANT: Carney, Darrell H.

;; TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
;; TITLE OF INVENTION: PEPTIDES
;; FILE REFERENCE: 3033.1000-008
;; CURRENT APPLICATION NUMBER: US/10/050,611
;; CURRENT FILING DATE: 2002-01-16
;; PRIOR APPLICATION NUMBER: 09/904,090
;; PRIOR FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: 60/217,583
;; PRIOR FILING DATE: 2000-07-12
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: human fragment of thrombin
US-10-050-611-3

Query Match 100.0%; Score 59; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
Db 14 CEGDSGGPFV 23

RESULT 9

US-10-050-611-4
;; Sequence 4, Application US/10050611
;; Publication No. US20020187933A1
;; GENERAL INFORMATION:
;; APPLICANT: Carney, Darrell H.
;; TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
;; TITLE OF INVENTION: PEPTIDES
;; FILE REFERENCE: 3033.1000-008
;; CURRENT APPLICATION NUMBER: US/10/050,611
;; CURRENT FILING DATE: 2002-01-16
;; PRIOR APPLICATION NUMBER: 09/904,090
;; PRIOR FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: 60/217,583
;; PRIOR FILING DATE: 2000-07-12
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: c-terminal amidated fragment of human thrombin
;; FEATURE:
;; NAME/KEY: AMIDATION
;; LOCATION: 23
;; OTHER INFORMATION: valine is amidated as CONH2
US-10-050-611-4

Query Match 100.0%; Score 59; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
Db 14 CEGDSGGPFV 23

RESULT 10

US-10-050-688-5
;; Sequence 5, Application US/10050688
;; Publication No. US20020198154A1
;; GENERAL INFORMATION:
;; APPLICANT: Carney, Darrell H.
;; APPLICANT: Crowther, Roger S.

APPLICANT: Stiernberg, Janet
APPLICANT: Bergmann, John
TITLE OF INVENTION: STIMULATION OF CARTILAGE GROWTH WITH
TITLE OF INVENTION: AGONISTS OF THE NON-PROTEOLYTICALLY ACTIVATED THROMBIN
TITLE OF INVENTION: RECEPTOR
FILE REFERENCE: 3033.1003-004
CURRENT APPLICATION NUMBER: US/10/050.688
CURRENT FILING DATE: 2002-01-16
PRIOR APPLICATION NUMBER: 09/909,348
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/219,800
PRIOR FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 5
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptide fragment of thrombin
US-10-050-688-5

Query Match 100.0%; Score 59; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
| | | | | | | |
Db 14 CEGDSGGPFV 23

RESULT 11.
US-10-050-688-6
Sequence 6, Application US/10050688
Publication No. US20020198154A1
GENERAL INFORMATION:
APPLICANT: Carney, Darrell H.
APPLICANT: Growther, Roger S.
APPLICANT: Stiernberg, Janet
APPLICANT: Bergmann, John
TITLE OF INVENTION: STIMULATION OF CARTILAGE GROWTH WITH
TITLE OF INVENTION: AGONISTS OF THE NON-PROTEOLYTICALLY ACTIVATED THROMBIN
FILE REFERENCE: 3033.1003-004
CURRENT APPLICATION NUMBER: US/10/050.688
CURRENT FILING DATE: 2002-01-16
PRIOR APPLICATION NUMBER: 09/909,348
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/219,800
PRIOR FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 6
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptide fragment of thrombin
NAME/KEY: AMIDATION
LOCATION: (23)...(23)
OTHER INFORMATION: CONH2
US-10-050-688-6

Query Match 100.0%; Score 59; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
| | | | | | | |
Db 14 CEGDSGGPFV 23

RESULT 12
US-09-777-328-8
Sequence 8, Application US/09777328
Patent No. US20020032314A1
GENERAL INFORMATION:
APPLICANT: CAREY, DARRELL H.
APPLICANT: RAMAKRISHNAN, SHYAM
TITLE OF INVENTION: SYNTHETIC PEPTIDE NEUTROPHIL CELL CHEMOTACTIC AGENTS
FILE REFERENCE: CHBP:002
CURRENT APPLICATION NUMBER: US/09/777,328
CURRENT FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 08/330,594
PRIOR FILING DATE: 1994-10-28
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-777-328-8

Query Match 100.0%; Score 59; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
| | | | | | | |
Db 14 CEGDSGGPFV 23

RESULT 13
US-09-904-090-3
Sequence 3, Application US/09904090
Patent No. US20020061852A1
GENERAL INFORMATION:
APPLICANT: Carney, Darrell
TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
FILE REFERENCE: 3033.1000-001
CURRENT APPLICATION NUMBER: US/09/904,090
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: US 60/217,583
PRIOR FILING DATE: 2000-07-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Polypeptide, fragment of thrombin
US-09-904-090-3

Query Match 100.0%; Score 59; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
| | | | | | | |
Db 14 CEGDSGGPFV 23

RESULT 14
US-10-050-692-5
Sequence 5, Application US/10050692
Publication No. US2002018205A1
GENERAL INFORMATION:
APPLICANT: Carney, Darrell H.
APPLICANT: Growther, Roger S.
APPLICANT: Simmons, David J.

;; APPLICANT: Yang, Jinping
;; APPLICANT: Rebin, William R.
;; TITLE OF INVENTION: STIMULATION OF BONE GROWTH WITH THROMBIN
;; TITLE OF INVENTION: PEPTIDE DERIVATIVES
;; FILE REFERENCE: 3033.1002-004
;; CURRENT APPLICATION NUMBER: US/10/050,692
;; CURRENT FILING DATE: 2002-01-16
;; PRIOR APPLICATION NUMBER: 09/909,122
;; PRIOR FILING DATE: 2001-07-19
;; PRIOR APPLICATION NUMBER: 60/219,300
;; PRIOR FILING DATE: 2000-07-19
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: FastSEQ for Windows Version 4.0
;; SEQ ID NO 5
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: fragment of human prothrombin
US-10-050-692-5

Query Match 100.0%; Score 59; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CEGDSGGPFV 10
Db 16 CEGDSGGPFV 25

RESULT 15
US-09-909-348-5
;; Sequence 5, Application US/09909348
;; Patent No. US20020042373A1
;; GENERAL INFORMATION:
;; APPLICANT: Carney, Darrell H.
;; APPLICANT: Crowther, Roger S.
;; APPLICANT: Stierberg, Janet
;; APPLICANT: Bergmann, John
;; TITLE OF INVENTION: Stimulation Of Cartilage Growth With Agonists
;; TITLE OF INVENTION: Of The No. US20020042373A1-Proteolytically Activated Thrombin Re
;; FILE REFERENCE: 3033.1003-001
;; CURRENT APPLICATION NUMBER: US/09/909,348
;; CURRENT FILING DATE: 2001-07-19
;; PRIOR APPLICATION NUMBER: US 60/219,800
;; PRIOR FILING DATE: 2000-07-20
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: FastSEQ for Windows Version 4.0
;; SEQ ID NO 5
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Peptide fragment of Thrombin
US-09-909-348-5

Query Match 100.0%; Score 59; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CEGDSGGPFV 10
Db 16 CEGDSGGPFV 25

Search completed: May 12, 2003, 15:59:08
Job time : 21.4 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:31:22 ; Search time 28.6 Seconds
(without alignments)
46.591 Million cell updates/sec

Title: US-09-909-348-1
Perfect score: 59
Sequence: 1 CEGDGGPFV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	59	100.0	10	AAE20155	Serine esterase co
2	59	100.0	10	AAU78372	Serine esterase co
3	59	100.0	12	AAW50857	Serine esterase co
4	59	100.0	23	AAW83414	Cell growth/adhesi
5	59	100.0	23	AA112893	Nerve tissue regen
6	59	100.0	23	AA170363	Human thrombin rec
7	59	100.0	23	AAE22563	Human thrombin hig
8	59	100.0	23	AAE20159	Human thrombin pep
9	59	100.0	23	AAU78376	Thrombin peptide d
10	59	100.0	23	AAW50858	Thrombin-derived p

11	59	100.0	111	20	AAW99113	Bovine zeta 2 pret
12	59	100.0	116	20	AAW99115	Human zeta 2 pret
13	59	100.0	259	18	AAW11545	Human thrombin Asn
14	59	100.0	295	16	AAW74775	Wild-type thrombin
15	59	100.0	295	16	AAW74776	Mutant thrombin K5
16	59	100.0	295	16	AAW74777	Mutant thrombin E2
17	59	100.0	295	16	AAW74778	Mutant thrombin E2
18	59	100.0	295	16	AAW74779	Mutant thrombin E2
19	59	100.0	295	16	AAW74780	Mutant thrombin E2
20	59	100.0	295	16	AAW76033	Mutant thrombin R2
21	59	100.0	295	16	AAW76034	Mutant thrombin R2
22	59	100.0	295	16	AAW76035	Mutant thrombin R2
23	59	100.0	295	16	AAW76036	Mutant thrombin R2
24	59	100.0	295	16	AAW76037	Mutant thrombin W5
25	59	100.0	295	16	AAW76038	Mutant thrombin K5
26	59	100.0	295	16	AAW76039	Mutant thrombin W5
27	59	100.0	295	16	AAW76040	Mutant thrombin W5
28	59	100.0	295	18	AAW22892	Human mature throm
29	59	100.0	295	21	AAW88633	Amino acid sequenc
30	59	100.0	308	20	AAW99107	Bovine prothrombin
31	59	100.0	308	20	AAW99109	Human prothrombin
32	59	100.0	376	14	AAW41197	CD4/Thrombin fusio
33	59	100.0	376	20	AAW42789	Human CD4-thrombin
34	59	100.0	376	23	AAU10703	Prothrombin (PT)
35	59	100.0	579	14	AAW35763	Human prothrombin
36	59	100.0	579	18	AAW11546	Human prothrombin
37	59	100.0	579	18	AAW11544	Human prothrombin
38	59	100.0	579	20	AAW99108	Bovine prothrombin
39	59	100.0	582	20	AAW99106	Human prothrombin
40	59	100.0	615	14	AAW38741	Human prothrombin
41	59	100.0	615	17	AAW36216	Human prothrombin
42	59	100.0	615	17	AAW90377	Human prothrombin
43	59	100.0	622	18	AAW11543	Human prothromb
44	59	100.0	622	20	AAW49566	Platelet membrane
45	56	94.9	164	22	AAW68931	Human protease dom

ALIGNMENTS

RESULT 1
AAE20155
ID AAE20155 standard; peptide; 10 AA.
XX
AC AAE20155;
XX
DT 18-JUN-2002 (first entry)
XX
DE Serine esterase conserved peptide #1.
XX
KW Cartilage growth; cartilage repair; arthritic joint; traumatic injury;
KW non-proteolytically activated thrombin receptor; NPAR; chondrocyte;
KW therapy; implantation; serine esterase conserved peptide.
XX
OS Unidentified.
XX
PN WO200207748-A2.
XX
PD 31-JAN-2002.
XX
PF 19-JUL-2001; 2001WO-US22668.
XX
PR 20-JUL-2000; 2000US-219800P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Carney DH, Crowther RS, Stiernberg J, Bergmann J;
XX
DR WPI; 2002-268953/31.
XX
PT Stimulating growth and repair of cartilage, useful for treating e.g.
PT arthritis, by local administration of an agonist of non-proteolytically
PT activated thrombin receptor

hand date

XX PS Claim 7; Page 24; 28pp; English.

CC The invention relates to a method of stimulating growth and repair of

CC cartilage. The method involves administering to the site, an agonist

CC of non-proteolytically activated thrombin receptor (NPAR). The method

CC is used in human or veterinary medicine for the treatment of arthritic

CC joints and damage/loss of cartilage caused by traumatic injury. Also

CC chondrocytes may be cultured in presence of NPAR agonist to provide

CC cells for implantation at sites requiring growth/repair of cartilage.

CC The present sequence is serine esterase conserved peptide. This

CC sequence is present in the thrombin peptide derivatives which serve

CC as a NPAR agonist.

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 59; DB 23; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.081;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10

Db | | | | | | | |

1 CEGDSGGPFV 10

RESULT 2

AAU78372

ID AAU78372 standard; Peptide; 10 AA.

AC AAU78372;

XX 18-JUN-2002 (first entry)

DT Serine esterase conserved sequence #1.

DE Serine esterase; thrombin; receptor; agonist;

KW bone growth stimulation; osteoinduction; farm animal; companion animal;

KW laboratory animal; bone graft; segmental bone gap; bone void;

KW non-union fracture.

XX Synthetic.

OS WO200205836-A2.

XX 24-JAN-2002.

XX 18-JUL-2001; 2001WO-US22641.

XX 19-JUL-2000; 2000US-219300P.

XX (TEXA) UNIV TEXAS SYSTEM.

PA Carney DH, Crowther RS, Simmons DJ, Yang J, Redin WR;

PI WPI; 2002-303796/34.

XX Stimulating bone growth at a site in a subject in need of

PT osteoinduction, such as a site of bone graft, segmental bone gap, bone

PT void or non-union structure, by administering agonist of activated

PT thrombin receptor.

XX Claim 6; Page 21; 27pp; English.

XX The invention describes a method of stimulating bone growth at a site

CC in a subject in need of osteoinduction. The method involves administering

CC an agonist to stimulate bone growth at a site in a subject (e.g. a farm

CC animal, companion animal or laboratory animal), in need of

CC osteoinduction, such as the site in need of a bone graft in a subject, a

CC segmental bone gap, a bone void or a non-union fracture. This sequence

CC represents a serine esterase conserved sequence obtained from a serine

CC esterase that can stimulate or activate the non-proteolytically

CC activated thrombin receptor.

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 59; DB 23; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.081;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10

Db | | | | | | | |

1 CEGDSGGPFV 10

RESULT 3

AAU50857

ID AAU50857 standard; Peptide; 12 AA.

XX AAU50857;

XX 01-MAY-2002 (first entry)

DT Serine esterase conserved sequence used in cardiac tissue repair.

DE Serine esterase; thrombin; revascularisation; vascular occlusion;

KW tissue repair; vulnery; vasotropic; cardiac; angiogenesis;

KW restenosis; therapy; enzyme; human.

XX Homo sapiens.

OS WO200204008-A2.

XX 17-JAN-2002.

XX 12-JUL-2001; 2001WO-US21944.

XX 12-JUL-2000; 2000US-217583P.

XX (TEXA) UNIV TEXAS SYSTEM.

PA Carney DH;

PI WPI; 2002-179665/23.

XX Promoting cardiac tissue repair, stimulating revascularisation,

PT stimulating vascular endothelial cell proliferation, and inhibiting

PT vascular occlusion by using angiogenic thrombin derivative peptide

XX Claim 3; Page 19; 24pp; English.

XX The present peptide comprises a thrombin-derived serine esterase

CC conserved sequence that is used in a claimed method for promoting

CC cardiac tissue repair. The method involves administering an

CC angiogenic thrombin-derived peptide, especially a thrombin receptor

CC binding domain comprising the 4-amino acid peptide given in

CC AAU50856 together with the serine esterase conserved sequence,

CC or preferably the peptide given in AAU50858, which includes both

CC these peptide sequences. The thrombin-derived peptide is

CC administered during or following cardiac surgery by injection

CC into cardiac tissue, and may be formulated as a sustained release

CC formulation. It is used in claimed methods of stimulating

CC revascularisation, stimulating vascular endothelial cell

CC proliferation, inhibiting vascular occlusion, and inhibiting

CC restenosis following balloon angioplasty, in which case the

CC peptide may be coated onto the catheter.

XX SQ Sequence 12 AA;

Query Match 100.0%; Score 59; DB 23; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.094;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10

Db | | | | | | | |

3 CEGDSGGPFV 12

~~1~~

3 X

23

1980-1981

STANDARDIZATION OF THE
TESTS AND MEASUREMENTS

XX The present invention describes a synthetic peptide (I) which is a
 CC neutrophil cell chemotactic agent. (I) has vulnerary and
 CC antiinflammatory activities. (I) is useful as a potent neutrophil cell
 CC chemotactic agent and for generating antibodies against the peptides,
 CC which are useful for modulating neutrophil recruitment to a wound site
 CC for enhancing or inhibiting inflammation and early effects of wound
 CC healing. Neutrophil response to (I) is specific, since monocytes and
 CC fibroblasts do not show any expression of the receptor to which (I)
 CC binds. The present sequence represents a human thrombin receptor binding
 CC domain peptide which is used in an example from the present invention.

XX SQ Sequence 23 AA;

Query Match 100.0%; Score 59; DB 22; Length 23;

Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
 Db 14 CEGDSGGPFV 23
 |||||

RESULT 7

AAE22563
 ID AAE22563 standard; peptide; 23 AA.

XX AC AAE22563;

XX DT 26-JUL-2002 (first entry)

XX DE Human thrombin high affinity receptor binding domain.

XX KW Human; proteolytically activated receptor for thrombin; neutrophil;
 KW chemotactic agent; PAR; inflammation; wound healing; chemotaxis;
 KW immune response; vulnerary; thrombin; receptor binding domain.

XX OS Homo sapiens.

XX PN US2002032314-A1.

XX PD 14-MAR-2002.

XX PF 05-FEB-2001; 2001US-0777328.

XX PR 28-OCT-1994; 94US-0330594.

XX PA (CHRY-) CHRYSLIS BIOTECHNOLOGY INC.

XX PI Carney DH, Ramakrishnan S;

XX DR WPI; 2002-371207/40.

XX PT New synthetic peptide neutrophil cell chemotactic agents, useful for
 PT stimulating or modulating neutrophil cell chemotactic migration,
 PT particularly for modulating neutrophil recruitment during immune
 PT response or in wound healing

XX PS Example 2; Page 3; 15pp; English.

XX The present invention relates to novel synthetic peptides and antibodies
 CC which are potent chemotactic agents for neutrophils. The peptides of the
 CC invention mimic the activity and role of the cleavage fragment of the
 CC proteolytically activated receptor for thrombin (PAR). They are useful
 CC for stimulating or modulating neutrophil cell chemotactic migration or
 CC for generating an antibody. In particular, the peptides of the invention
 CC are useful for modulating neutrophil recruitment to a wound site for
 CC enhancing or inhibiting inflammation and early effects in wound healing.
 CC They are also useful for modulated neutrophil chemotaxis in immune
 CC response. The present sequence is high affinity receptor binding
 CC domain of human thrombin. This peptide is used in the exemplification
 CC of the invention.

XX

SQ Sequence 23 AA;

Query Match 100.0%; Score 59; DB 23; Length 23;

Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
 Db 14 CEGDSGGPFV 23
 |||||

RESULT 8

AAE20159
 ID AAE20159 standard; peptide; 23 AA.

XX AC AAE20159;

XX DT 18-JUN-2002 (first entry)

XX DE Human thrombin peptide derivative #2.

XX KW Cartilage growth; cartilage repair; arthritic joint; traumatic injury;
 KW non-proteolytically activated thrombin receptor; NPAR; chondrocyte;
 KW therapy; implantation; thrombin peptide; human.

XX OS Homo sapiens.

XX PN WO200207748-A2.

XX PD 31-JAN-2002.

XX PF 19-JUL-2001; 2001WO-US22668.

XX PR 20-JUL-2000; 2000US-219800P.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PI Carney DH, Crowther RS, Stiernberg J, Bergmann J;

XX DR WPI; 2002-268953/31.

XX PT Stimulating growth and repair of cartilage, useful for treating e.g.
 PT arthritis, by local administration of an agonist of non-proteolytically
 PT activated thrombin receptor

XX PS Claim 12; Page 25; 28pp; English.

XX The invention relates to a method of stimulating growth and repair of
 CC cartilage. The method involves administering to the site, an agonist
 CC of non-proteolytically activated thrombin receptor (NPAR). The method
 CC is used in human or veterinary medicine for the treatment of arthritic
 CC joints and damage/loss of cartilage caused by traumatic injury. Also
 CC chondrocytes may be cultured in presence of NPAR agonist to provide
 CC cells for implantation at sites requiring growth/repair of cartilage.
 CC The present sequence is human thrombin peptide derivative which serves
 CC as a NPAR agonist.

XX SQ Sequence 23 AA;

Query Match 100.0%; Score 59; DB 23; Length 23;

Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
 Db 14 CEGDSGGPFV 23
 |||||

RESULT 9

AAU78376
 ID AAU78376 standard; Peptide; 23 AA.

XX AC AAU78376;

XX DT 18-JUN-2002 (first entry)
 XX DE Thrombin peptide derivative TP508.
 XX DE Thrombin; osteopathic; receptor; agonist; bone growth stimulation;
 KW osteoinduction; farm animal; companion animal; laboratory animal;
 KW bone graft; segmental bone gap; bone void; non-union fracture.
 XX OS Synthetic.
 XX XX
 XX XX
 FH Key Location/Qualifiers
 FT Misc-difference 3 /label= Unknown
 FT FT
 XX WO200205836-A2.
 XX XX
 PD 24-JAN-2002.
 XX XX
 XX 18-JUL-2001; 2001WO-US22641.
 XX XX
 PR 19-JUL-2000; 2000US-219300P.
 XX XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX Carney DH, Crowther RS, Simmons DJ, Yang J, Redin WR;
 PI WPI; 2002-303796/34.
 DR XX
 XX Stimulating bone growth at a site in a subject in need of
 FT osteoinduction, such as a site of bone graft, segmental bone gap, bone
 FT void or non-union structure, by administering agonist of activated
 FT thrombin receptor
 XX XX
 PS Claim 11; Page 22; 27pp; English.
 CC The invention describes a method of stimulating bone growth at a site
 CC in a subject in need of osteoinduction. The method involves administering
 CC an agonist to stimulate bone growth at a site in a subject (e.g. a farm
 CC animal, companion animal or laboratory animal), in need of
 CC osteoinduction, such as the site in need of a bone graft in a subject, a
 CC segmental bone gap, a bone void or a non-union fracture. This sequence
 CC represents a thrombin peptide derivative obtained from a serine esterase
 CC that can stimulate or activate the non-proteolytically activated thrombin
 CC receptor.
 XX XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 59; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEGDSGGPFV 10
 DB 14 CEGDSGGPFV 23
 RESULT 10
 AAM50858
 ID AAM50858 standard; Peptide; 23 AA.
 XX XX
 AC AAM50858;
 XX XX
 DT 01-MAY-2002 (first entry)
 DE Thrombin-derived peptide used to promote cardiac tissue repair.
 XX XX
 KW Thrombin; revascularisation; vascular occlusion; tissue repair;
 KW vulnery; vasotropic; cardiant; angiogenesis; restenosis;
 KW therapy; human.
 XX XX
 OS Homo sapiens.
 XX XX

FH Key Location/Qualifiers
 FT Peptide 10..13 "thrombin receptor binding domain"
 FT Peptide 12..23
 FT /note- "serine esterase conserved sequence"
 XX WO200204008-A2.
 XX 17-JAN-2002.
 XX XX
 XX 12-JUL-2001; 2001WO-US21944.
 XX XX
 PR 12-JUL-2000; 2000US-217583P.
 XX XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX Carney DH;
 PI WPI; 2002-179665/23.
 DR XX
 XX Promoting cardiac tissue repair, stimulating revascularisation,
 FT stimulating vascular endothelial cell proliferation, and inhibiting
 FT vascular occlusion by using angiogenic thrombin derivative peptide
 XX XX
 PS Claim 4; Page 19; 24pp; English.
 CC The present peptide comprises a thrombin-derived peptide, TP508,
 CC that includes a thrombin receptor binding domain sequence (see also
 CC AAM50856) and a serine esterase conserved sequence (see also
 CC AAM50857). The peptide is used in a claimed method for promoting
 CC cardiac tissue repair. It is administered during or following
 CC cardiac surgery by injection into cardiac tissue, and may be
 CC formulated as a sustained release formulation. The thrombin
 CC derivative peptide is also used in claimed methods of stimulating
 CC revascularisation, stimulating vascular endothelial cell
 CC proliferation, inhibiting vascular occlusion, and inhibiting
 CC restenosis following balloon angioplasty, in which case it may be
 CC coated onto the catheter.
 XX XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 59; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEGDSGGPFV 10
 DB 14 CEGDSGGPFV 23
 RESULT 11
 AAM99113
 ID AAM99113 standard; protein; 111 AA.
 XX XX
 AC AAM99113;
 XX XX
 DT 14-MAY-1999 (first entry)
 DE Bovine zeta 2 prethrombin 2.
 XX XX
 KW Prothrombin; exosite assay; anticoagulant; blood clot; thrombin;
 KW cardiovascular disease; stroke; haematological disorder.
 XX XX
 OS Bos sp.
 XX XX
 PN WO9855130-A1.
 XX 10-DEC-1998.
 XX XX
 PF 28-MAY-1998; 98WO-US10840.
 XX XX
 PR 08-APR-1998; 98US-0081030.
 PR 06-JUN-1997; 97US-0048864.

XX PA (UYEM-) UNIV EMORY.
 XX PT Krishnaswamy S;
 XX PT WPI; 1999-070237/06.
 XX DR
 XX PS Disclosure; Page 42-43; 61pp; English.
 XX CC An exosite assay has been developed for inhibition of the catalytic
 XX CC cleavage of prothrombin (Pth) to thrombin (Th) by prothrombinase (I), at
 XX CC a site remote from the catalytic site of (I) comprises: (a) preparing a
 XX CC solution containing 0.05-20 mu M substrate (S), that includes a protease
 XX CC cleavage site and exosite-binding determinant; 0.05-200 nM factor Va;
 XX CC 30-500 micro M phospholipids (PL); test inhibitor (A) in buffer of pH
 XX CC 7-9, containing 1-10 mM calcium ions but no calcium-chelating agent;
 XX CC (b) initiating catalytic cleavage of (S) by adding an aliquot of factor
 XX CC Xa to final concentration 0.05-200 nM so that there is an excess of Va
 XX CC over Xa, forming a S/(I) complex; (c) withdrawing aliquots of the
 XX CC reaction mixture, quenching them; and (d) assaying for concentration of
 XX CC Th. Alternatively, in the initial solution S is replaced by the same
 XX CC concentration of Xa (less than the amount of Va), and reaction is started
 XX CC by adding S. Also described in the present invention are inhibitors (A') are
 XX CC having IC50 less than 1 mu M identified by this assay. (A') are
 XX CC potentially useful as a new class of anticoagulants for treatment of
 XX CC cardiovascular disease, stroke and haematological disorders. The method
 XX CC is based on the finding that exosite interactions are critical for
 XX CC substrate specificity in catalytic formation of Th. The present sequence
 XX CC represents bovine zeta 2 prethrombin 2.
 XX SQ Sequence 111 AA;
 Query Match 100.0%; Score 59; DB 20; Length 111;
 Best Local Similarity 100.0%; Pred. No. 0.6;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEGDSGGPFV 10
 Db 53 CEGDSGGPFV 62
 RESULT 12
 AAW91115
 ID AAW91115 standard; protein; 116 AA.
 XX AC AAW91115;
 XX DT 14-MAY-1999 (first entry)
 XX DE Human zeta 2 prethrombin 2.
 XX KW Prothrombin; exosite assay; anticoagulant; blood clot; thrombin;
 XX KW cardiovascular disease; stroke; haematological disorder.
 XX OS Homo sapiens.
 XX PN WO9855130-A1.
 XX PD 10-DEC-1998.
 XX PF 28-MAY-1998; 98WO-US10840.
 XX PR 08-APR-1998; 98US-0081030.
 XX PR 06-JUN-1997; 97US-0048864.
 XX PA (UYEM-) UNIV EMORY.
 XX PI Krishnaswamy S;
 XX XX

XX PA (UYEM-) UNIV EMORY.
 XX PT Krishnaswamy S;
 XX PT WPI; 1999-070237/06.
 XX DR
 XX PS Disclosure; Page 44-45; 61pp; English.
 XX CC An exosite assay has been developed for inhibition of the catalytic
 XX CC cleavage of prothrombin (Pth) to thrombin (Th) by prothrombinase (I), at
 XX CC a site remote from the catalytic site of (I) comprises: (a) preparing a
 XX CC solution containing 0.05-20 mu M substrate (S), that includes a protease
 XX CC cleavage site and exosite-binding determinant; 0.05-200 nM factor Va;
 XX CC 30-500 micro M phospholipids (PL); test inhibitor (A) in buffer of pH
 XX CC 7-9, containing 1-10 mM calcium ions but no calcium-chelating agent;
 XX CC (b) initiating catalytic cleavage of (S) by adding an aliquot of factor
 XX CC Xa to final concentration 0.05-200 nM so that there is an excess of Va
 XX CC over Xa, forming a S/(I) complex; (c) withdrawing aliquots of the
 XX CC reaction mixture, quenching them; and (d) assaying for concentration of
 XX CC Th. Alternatively, in the initial solution S is replaced by the same
 XX CC concentration of Xa (less than the amount of Va), and reaction is started
 XX CC by adding S. Also described in the present invention are inhibitors (A') are
 XX CC having IC50 less than 1 mu M identified by this assay. (A') are
 XX CC potentially useful as a new class of anticoagulants for treatment of
 XX CC cardiovascular disease, stroke and haematological disorders. The method
 XX CC is based on the finding that exosite interactions are critical for
 XX CC substrate specificity in catalytic formation of Th. The present sequence
 XX CC represents human zeta.2 prethrombin 2.
 XX SQ Sequence 116 AA;
 Query Match 100.0%; Score 59; DB 20; Length 116;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEGDSGGPFV 10
 Db 58 CEGDSGGPFV 67
 RESULT 13
 AAW11545
 ID AAW11545 standard; protein; 259 AA.
 XX AC AAW11545;
 XX DT 01-OCT-1997 (first entry)
 XX DE Human thrombin Asn99 mutant.
 XX KW Prothrombin; mutant; muten; platelet aggregation; blood clotting;
 XX KW coagulation; reduce; decrease; hirudin; heparin; anti-thrombin III;
 XX KW antagonist; D99N.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 XX FT Protein 1..259
 XX FT /label= thrombin_Asn99
 XX FT Misc-difference 99
 XX FT /note= "Wild-type Asp residue has been replaced by
 XX FT Asn"
 XX XX WO9641868-A2.
 XX PN 27-DEC-1996.
 XX PD 12-JUN-1996; 96WO-AF00105.
 XX PR 13-JUN-1995; 95AT-0001006.
 XX PA (IMMO) IMMUNO AG.

XX Eibl J, Falkner F, Fischer B, Mitterer A, Schlokat U;
 XX WPI; 1997-065455/06.
 XX Prothrombin mutants with reduced clotting activity - useful as
 PT antagonists of thrombin inhibitors or for anticoagulant therapy
 XX
 PS Example 3; Page -: 73pp; German.

XX Prothrombin mutants having one or more changes in amino acid sequence
 CC compared with the natural protein and having 0-10% (preferably 0-0.25%)
 CC of the activity of the natural protein are claimed, provided that the
 CC changes in amino acid sequence do not affect the capacity of the
 CC mutants to bind to specific ligands and receptors. The mutants have
 CC greatly reduced clotting activity and are useful as antagonists of
 CC thrombin inhibitors such as hirudin, heparin and anti-thrombin III.
 CC The mutations may also result in changes to the in vivo half-life
 CC of prothrombin. The half-life may be reduced to less than 10 minutes
 CC or the mutant prothrombin may have an extended half-life of more than
 CC 1 hour, making it useful as an anticoagulant and to inhibit side-
 CC effects of anti-coagulant treatment. They are converted to inactive
 CC thrombin and are able to compete with native, active thrombin for
 CC binding to receptors. The present sequence represents the thrombin
 CC mutant which is derived by trypsin cleavage of a specifically
 CC claimed human prothrombin mutant in which Asp at position 419 is
 CC changed to Asn. The thrombin Asn99 mutant was found to have only
 CC 0.24% of the activity of wild-type thrombin on a chromogenic
 CC substrate.

CC (Note: This sequence does not appear in the specification and has
 CC been produced by modifying the wild-type sequence of human
 CC prothrombin which appears in figure 1).

XX SQ Sequence 259 AA;

Query Match 100.0%; Score 59; DB 18; Length 259;
 Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0;

QY 1 CEGDSGGPFV 10
 |||||
 DB 201 CEGDSGGPFV 210

RESULT 14
 AAR74775
 ID AAR74775 standard; Protein; 295 AA.

XX AAR74775;
 XX 04-NOV-1995 (first entry)

DE Wild-type thrombin.

KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;
 KW anticoagulant; protein engineering; ss.

XX Homo sapiens.

PH Key Location/Qualifiers
 FT Protein 37..295
 FT /note= "mature protein"

XX WO9513385-A.

PD 18-MAY-1995.

XX 14-NOV-1994; 94WO-US13104.

XX 10-JUN-1994; 94US-0258038.

XX 12-NOV-1993; 93US-0152657.

XX (GILE-) GILEAD SCI.

XX Gibbs CS, Leung LLK, Tsiang M;
 XX WPI; 1995-194103/25.
 DR N-PSDB; AAQ92455.
 XX Thrombin derivs with segregated pro- and anticoagulant activities
 PT useful for treating thrombotic disorders but also diagnosis,
 PT treatment of tumours, etc.

XX Disclosure; Fig 1; 78pp; English.

XX The sequence represents wild-type (reference) thrombin. Mutants
 CC of this sequence (AAR74776-80 and AAR76033-41) have at least 80%
 CC homology with thrombin, and are capable of protein-C activation
 CC without significant fibrinogen clotting activity, and vice versa
 CC (specifically have a ratio of protein-C activity to fibrinogen
 CC clotting activity of less than 0.5 or greater than 2 compared to
 CC thrombin). The mutant thrombin sequences, produced in recombinant
 CC cell culture or by in vitro methods, and are used to treat
 CC thrombotic conditions, particularly during cardiac bypass surgery
 CC and in cases of septic shock.

XX SQ Sequence 295 AA;

Query Match 100.0%; Score 59; DB 16; Length 295;
 Best Local Similarity 100.0%; Pred. No. 1.4; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0;

QY 1 CEGDSGGPFV 10
 |||||
 DB 237 CEGDSGGPFV 246

RESULT 15

AAR74776
 ID AAR74776 standard; Protein; 295 AA.

XX AAR74776;

XX 04-NOV-1995 (first entry)

XX Mutant thrombin K52A, R233A.

XX Thrombin; oligonucleotide-directed mutagenesis; procoagulant;
 KW anticoagulant; protein engineering; ss.

XX Homo sapiens.

PH Key Location/Qualifiers
 FT Misc-difference 88
 FT /note= "Lys in wild-type"
 FT Misc-difference 269
 FT /note= "Arg in wild-type"
 FT Protein 37..295
 FT /note= "mature protein"

XX WO9513385-A.

PD 18-MAY-1995.

XX 14-NOV-1994; 94WO-US13104.

XX 10-JUN-1994; 94US-0258038.

XX 12-NOV-1993; 93US-0152657.

XX (GILE-) GILEAD SCI.

XX Gibbs CS, Leung LLK, Tsiang M;

XX WPI; 1995-194103/25.

XX Thrombin derivs with segregated pro- and anticoagulant activities

PT useful for treating thrombotic disorders but also diagnosis,
PT treatment of tumours, etc.

XX

XX PS Claim 22; Page 63/3; 78pp; English.

XX

CC The mutant thrombin sequence, generated by oligonucleotide-directed
CC mutagenesis, has at least 80% homology with thrombin and is
CC capable of protein-C activation without significant fibrinogen
CC clotting activity, and vice versa. (specifically, it has a ratio
CC of protein-C activity to fibrinogen clotting activity of less than
CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin
CC is produced in recombinant cell culture or by in vitro methods,
CC and is used to treat thrombotic conditions, particularly during
CC cardiac bypass surgery and in cases of septic shock.

XX

SO Sequence 295 AA;

Query Match

Best Local Similarity 100.0%; Score 59; DB 16; Length 295;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10

|||||

Db 237 CEGDSGGPFV 246

Search completed: May 12, 2003, 15:36:02
Job time : 29.6 secs

Result No.	Score	Query			DB	ID	Description
		Match	Length	DB			
1	59	100.0	234	13	Q90244	Q90244 acipenser t	
2	59	100.0	235	6	Q28731	Q28731 ocyotolagus	
3	59	100.0	235	13	Q91004	Q91004 gecko gecko	
4	59	100.0	235	13	Q90387	Q90387 cynops pyrr	
5	59	100.0	239	13	Q91218	Q91218 oncorhynch	
6	59	100.0	420	13	Q90504	Q90504 eptatretus	
7	59	100.0	607	13	Q91001	Q91001 gallus gall	
8	59	100.0	608	13	Q9PTW7	Q9PTW7 struthio ca	
9	56	94.9	255	5	Q9NBC9	Q9NBC9 glossina mo	
10	56	94.9	267	5	Q9BK47	Q9BK47 ludia foli	
11	56	94.9	202	11	Q9CW97	Q9CW97 mus musculus	
12	56	94.9	1524	13	Q91674	Q91674 xenopus lae	
13	53	89.8	250	5	Q9V514	Q9V514 drosophila	
14	53	89.8	252	5	Q76498	Q76498 diaphres a	
15	53	89.8	255	4	Q96RQ0	Q96RQ0 homo sapien	
16	53	89.8	358	5	Q45029	Q45029 drosophila	

SQ SEQUENCE 234 AA; 26846 MW; 45C558D6618E0585 CRC64;


```
Query Match      100.0%; Score 59; DB 13; Length 234;
Best Local Similarity 100.0%; Pred. No. 0.0097;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 177 CEGDSGGPFV 186

RESULT 2
Q28731 ID Q28731 PRELIMINARY; PRT; 235 AA.
AC Q28731
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Thrombin (Fragment).
GN THROMBIN.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., Macgillivray R.T.A.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT Amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
QY 1 CEGDSGGPFV 10
DB 177 CEGDSGGPFV 186

Query Match      100.0%; Score 59; DB 6; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 178 CEGDSGGPFV 187

RESULT 3
Q91004 ID Q91004 PRELIMINARY; PRT; 235 AA.
AC Q91004
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Thrombin (Fragment).
GN THROMBIN.
OS Gecko gekko (Tokay gekko).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae; Gekko.
OX NCBI_TaxID=36310;
RN [1]
RN SEQUENCE FROM N.A.
```

```
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., Macgillivray R.T.A.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT Amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
QY 1 CEGDSGGPFV 10
DB 177 CEGDSGGPFV 186

Query Match      100.0%; Score 59; DB 13; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 177 CEGDSGGPFV 186

RESULT 4
Q90387 ID Q90387 PRELIMINARY; PRT; 235 AA.
AC Q90387
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Thrombin (Fragment).
GN THROMBIN.
OS Cynops pyrrhogaster (Japanese common newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae; Cynops.
OX NCBI_TaxID=8330;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., Macgillivray R.T.A.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT Amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
QY 1 CEGDSGGPFV 10
DB 177 CEGDSGGPFV 186

Query Match      100.0%; Score 59; DB 13; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 177 CEGDSGGPFV 186

Query Match      100.0%; Score 59; DB 13; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 177 CEGDSGGPFV 186
```

```

Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 177 CEGDSGGPFV 186

RESULT 5
Q91218 ID Q91218 PRELIMINARY; PRT; 239 AA.
AC Q91218;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Thrombin (Fragment).
GN THROMBIN.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., MacGillivray R.T.A.;
RT "Partial Characterization of vertebrate prothrombin cDNAs:
RT Amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; M81398; AAA49433.1; -.
DR HSP; P00734; 187X.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
FT NON_TER 1
SQ SEQUENCE 239 AA; 27396 MW; F0F43F9A3205BF38 CRC64;

Query Match 100.0%; Score 59; DB 13; Length 239;
Best Local Similarity 100.0%; Pred. No. 0.0099;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 177 CEGDSGGPFV 186

RESULT 6
Q90504 ID Q90504 PRELIMINARY; PRT; 420 AA.
AC Q90504;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Thrombin.
OS Eptatretus stoutii (Pacific hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Eptatretinae; Eptatretus.
OX NCBI_TaxID=7765;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=1557383;

```

```

RA Banfield D.K., MacGillivray R.T.;
RT "Partial Characterization of vertebrate prothrombin cDNAs:
RT Amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=94223694; PubMed=7513365;
RA Banfield D.K., Irwin D.M., Walz D.A., MacGillivray R.T.;
RT "Evolution of prothrombin: Isolation and characterization of the cDNAs
RT encoding chicken and hagfish prothrombin.";
RL J. Mol. Evol. 38:177-187(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Banfield D.K.;
RT Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; M81393; AAA21620.1; -.
DR HSP; P00734; 10VS.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 420 AA; 47888 MW; 64522AA21A57B67A CRC64;

Query Match 100.0%; Score 59; DB 13; Length 420;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
DB 361 CEGDSGGPFV 370

RESULT 7
Q91001 ID Q91001 PRELIMINARY; PRT; 607 AA.
AC Q91001;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Thrombin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., MacGillivray R.T.;
RT "Partial Characterization of vertebrate prothrombin cDNAs:
RT Amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RN [2]

```

DR	EMBL; AB028871; BAA89046.1; -
DR	HSSP; P00734; IUVS.
DR	MEROPS; S01.217; -
DR	InterPro; IPRO01314; Chymotrypsin.
DR	InterPro; IPRO02383; GLA_blood.
DR	InterPro; IPRO00001; Kringle.
DR	InterPro; IPRO01254; Ser.protease_Try.
DR	InterPro; IPRO00294; VitK_dep_GLA.
DR	Pfam; PF00594; gla; 1.
DR	Pfam; PF00051; kringle; 2.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	PRINTS; PR00001; GLABLOOD.
DR	PRINTS; PR00018; KRINGLE.
DR	ProDom; PD000395; Kringle; 2.
DR	SMART; SM00069; GLA; 1.
DR	SMART; SM00130; KR; 2.
DR	SMART; SM00020; TRY_SPC; 1.
DR	PROSITE; PS00011; GLU CARBOXYLATION; 1.
DR	PROSITE; PS00021; KRINGLE_1; 2.
DR	PROSITE; PS00070; KRINGLE_2; 2.
DR	PROSITE; PS00240; TRYPsin_DOM; 1.
DR	PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR	PROSITE; PS00135; TRYPsin_SER; 1.
KW	Hydrolase; Serine protease.
SQ	SEQUENCE 508 AA; 69392 MW; 11B974B9AE54EA2 CRC64;
	Query Match 100.0%; Score 59; DB 13; Length 608;
	Best Local Similarity 100.0%; Pred. No. 0.025;
	Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps
QY	1 CEGDGGGPFV 10
Db	550 CEGDGGGPFV 559
RESULT 9	
Q9NBC9	
ID	Q9NBC9 PRELIMINARY; PRT; 255 AA.
AC	Q9NBC9;
DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Trypsin-like serine protease precursor.
OS	Glossina morsitans morsitans.
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Hippoboscidae; Glossiniidae; Glossina.
OX	NCBI_Taxid=37546; [1]
RN	SEQUENCE FROM N.A.
RC	TISSUE=GUT;
RP	MEDLINE=21135159; PubMed=11240636;
RX	Yan J., Cheng Q., Li C.B., Aksoy S.;
RA	"Molecular characterization of two serine proteases expressed in gut tissue of the African trypanosome vector, Glossina morsitans morsitans".
RT	Insect Mol. Biol. 10:47-56(2001).
RL	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPsin FAMILY.
CC	EMBL; AF252869; AAF91346.1; -
DR	HSSP; P00763; LDPO.
DR	MEROPS; S01.114; -
DR	InterPro; IPRO01314; Chymotrypsin.
DR	InterPro; IPRO01254; Ser.protease_Try.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	SMART; SM00020; TRY_SPC; 1.
DR	PROSITE; PS00240; TRYPsin_DOM; 1.
DR	PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR	PROSITE; PS00135; TRYPsin_SER; 1.
KW	Hydrolase; protease; Serine protease; Signal.
FT	POTENTIAL 1 19

FT CHAIN 28 255 TRYPsin-LIKE SERINE PROTEASE.
 SQ SEQUENCE 255 AA; 27565 MW; E914E31FE7DD6579 CRC64;
 Query Match 94.9%; Score 56; DB 5; Length 255;
 Best Local Similarity 90.0%; Pred. No. 0.036;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEGDSGGPFV 10
 Db 205 CQDGGGPFV 214
 I:|||||
 RESULT 10
 ID Q9BK47 PRELIMINARY; PRT; 267 AA.
 AC Q9BK47;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Sea STAR regeneration-associated protease SRAP.
 OS Luidia foliolata.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Asterozoa; Valvatacea; Paxillosida; Luidiidae; Luidia.
 OX NCBI_TaxID=105861;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21100442; PubMed=11179669;
 DR HSSP; P00763; IDPO.
 DR EMBL; AF312826; NAK15274.1; --
 DR MEROPS; S01.224; --
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser. protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS02440; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 267 AA; 28761 MW; 88F61A061921860C CRC64;

Query Match 94.9%; Score 56; DB 5; Length 267;
 Best Local Similarity 90.0%; Pred. No. 0.038;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
 Db 212 CQDGGGPFV 221
 I:|||||
 RESULT 11
 ID Q9CW97 PRELIMINARY; PRT; 502 AA.
 AC Q9CW97;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Adult male kidney cDNA, RIKEN full-length enriched library,
 clone:0610030A17, full insert sequence (Fragment).
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC EMBL; AK002694; BAB22289.1; --
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.224; --
 DR MGD; MGI:1196620; Hpn.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser. protease_Try.
 DR InterPro; IPR001190; Srcr_receptor.
 DR InterPro; IPR002155; Thiolase.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00202; SR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00098; THIOLEASE_1; UNKNOWN_1.
 DR PROSITE; PS02440; TRYPSIN_DOM; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 FT NON_TER 1
 SQ SEQUENCE 502 AA; 53069 MW; 97BE638938CC4FCC CRC64;
 Query Match 94.9%; Score 56; DB 11; Length 502;
 Best Local Similarity 90.0%; Pred. No. 0.072;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEGDSGGPFV 10
 Db 434 CQDGGGPFV 443
 I:|||||
 RESULT 12
 ID Q91674 PRELIMINARY; PRT; 1524 AA.
 AC Q91674;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Polyprotein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90432219; PubMed=10500163;
 RA Lindsay L.B., Yang J.C., Hedrick J.L.;
 RT "Ovocymase, a Xenopus laevis egg extracellular protease, is
 RT translated as part of an unusual polyprotease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:11253-11258(1999).
 RN [2]

RP SEQUENCE FROM N.A.
 RA Yang J.C., Lindsey L.L., Hedrick J.L.;
 RT "cDNA Cloning of Ovocysteine, a Chymotrypsin-like Protease Released
 RL From Xenopus laevis Eggs at Fertilization.";
 CC Submitted (MAR1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 4 CUB DOMAINS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; U01290; AAC24717.1; -;
 DR HSP; P00763; IDPO.
 DR MEROPS; S01.022; -;
 DR MEROPS; S01.245; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000859; CUB_domain.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00431; CUB; 5.
 DR Pfam; PF00089; trypsin; 3.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00042; CUB; 4.
 DR SMART; SM00020; TRYPSIN_DOM; 3.
 DR PROSITE; PS01180; CUB; 5.
 DR PROSITE; PS00240; TRYPSIN_DOM; 3.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_3.
 DR PROSITE; PS00135; TRYPSIN_SER; 3.
 KW Hydrolase; Serine protease.
 FT CHAIN 57 308 SERINE PROTEASE.
 FT CHAIN 584 817 SERINE PROTEASE.
 FT CHAIN 1295 1524 OVOCHYMASE.
 SQ SEQUENCE 1524 AA; 167566 MW; 32FE42128F37268 CRC64;

Query Match 94.9%; Score 56; DB 13; Length 1524;
 Best Local Similarity 90.0%; Pred. No. 0.22;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
 I:|||||||
 Db 243 CQDSDGGPFV 252

RESULT 13
 Q9V514 PRELIMINARY; PRT; 250 AA.
 ID Q9V514;
 AC Q9V514;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE CG11824 protein.
 GN CG11824.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananthan R.A., Lewis S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobarri C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AE003835; AAF59008.1; -;
 DR HSP; P00763; IDPO.
 DR Flybase; FBgn0033360; CG11824.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYPSIN_DOM; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 250 AA; 26328 MW; 1F54601A83D2F664 CRC64;

Query Match 89.8%; Score 53; DB 5; Length 250;
 Best Local Similarity 90.0%; Pred. No. 0.12;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
 I:|||||||
 Db 195 CEGDSGGPMV 204

RESULT 14
 O76498 PRELIMINARY; PRT; 252 AA.
 ID O76498;
 AC O76498;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Trypsin precursor.
 OS Diaprepes abbreviatus (Sugarcane rootstalk borer weevil).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Cucujiformia; Phytophaga; Curculionidae; Entiminae; Entimini;
 OC Diaprepes.
 OX NCBI_TaxID=13040;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99339928; PubMed=10411621;
 RA van X.-H., De Bondt H.L., Powell C.C., Bullock R.C., Borovsky D.,
 RT "Sequencing and characterization of the citrus weevil, Diaprepes
 RT abbreviatus, trypsin cDNA. Effect of aedes trypsin modulating oostatic
 RT factor on trypsin biosynthesis.";
 RL Eur. J. Biochem. 262:627-636(1999).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AF072690; AAC25596.1; -;
 DR HSP; P00763; IDPO.

DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal.
FT SIGNAL 1 11 POTENTIAL.
FT CHAIN 23 252 TRYPSIN.
SQ SEQUENCE 252 AA; 26064 MW; EEOEDBF116B042AB CRC64;

Query Match 89.8%; Score 53; DB 5; Length 252;
Best Local Similarity 80.0%; Pred. No. 0.12;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
Db 203 CGDSGGPFV 212

RESULT 15

Q96RQ0 PRELIMINARY; PRT; 255 AA.
AC Q96RQ0;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Prostinogen.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID-9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21226193; PubMed-11327827;
RA Takayama T.K., Carter C.A., Deng T.;
RT "Activation of prostate-specific antigen precursor (pro-PSA) by
RT prostin, a novel human prostatic serine protease identified by
RT degenerate PCR."
RL Biochemistry 40:1679-1687(2001).
DR EMBL; AF303046; AAK62813.1; -;
DR MEROPS; S01.081; -;
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 255 AA; 27986 MW; 00D5B79E14B9468F CRC64;

Query Match 89.8%; Score 53; DB 4; Length 255;
Best Local Similarity 90.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEGDSGGPFV 10
Db 204 CEGDSGGPLV 213

Search completed: May 12, 2003, 15:38:38
Job time : 24.6 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:32:07 ; Search time 5.8 Seconds
(without alignments)
71.511 Million cell updates/sec

Title: US-09-909-348-1
Perfect score: 59
Sequence: 1 CEGDSGPPV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	617	1	THRB_RAT
2	59	100.0	618	1	THRB_MOUSE
3	59	100.0	622	1	THRB_HUMAN
4	59	100.0	625	1	THRB_BOVIN
5	56	94.9	416	1	HEPS_MOUSE
6	56	94.9	417	1	HEPS_HUMAN
7	53	89.8	157	1	PTTC_CANE
8	53	89.8	157	1	PTTC_CAPHI
9	53	89.8	157	1	PTTC_FELCA
10	53	89.8	157	1	PTTC_HORSE
11	53	89.8	161	1	PTTC_MACMU
12	53	89.8	256	1	KLK_HUMAN
13	53	89.8	277	1	KLK_HUMAN
14	53	89.8	456	1	PTTC_BOVIN
15	53	89.8	459	1	PTTC_PIG
16	53	89.8	461	1	PTTC_HUMAN
17	52	88.1	232	1	VSP1_BOTJA
18	52	88.1	259	1	VSP1_VIPLE
19	52	88.1	275	1	FA9_RABIT
20	52	88.1	282	1	FA9_RAT
21	52	88.1	459	1	FA9_MOUSE
22	51	86.4	260	1	COGS_HYPLI
23	50	84.7	165	1	TRY3_LUCCU
24	50	84.7	218	1	CTR2_VESCR
25	50	84.7	227	1	TRYP_SACER
26	50	84.7	235	1	TRYD_HUMAN
27	50	84.7	238	1	TRY5_AEDAE
28	50	84.7	248	1	KLK6_HUMAN
29	50	84.7	248	1	KLK6_HUMAN
30	50	84.7	248	1	TRYP_FUSOX
31	50	84.7	250	1	KLK9_HUMAN
32	50	84.7	250	1	KLK8_HUMAN
33	50	84.7	251	1	KLKE_HUMAN

34 50 84.7 253 1 TRYB_DROER
35 50 84.7 253 1 TRYB_DROER
36 50 84.7 253 1 TRYD_DROER
37 50 84.7 253 1 TRYD_DROER
38 50 84.7 253 1 TRYD_DROER
39 50 84.7 254 1 TRY3_AEDAE
40 50 84.7 254 1 TRY3_SABU
41 50 84.7 255 1 TRY4_LUCCU
42 50 84.7 256 1 HYPA_HYPLI
43 50 84.7 256 1 HYPA_HYPLI
44 50 84.7 256 1 TRYD_DROER
45 50 84.7 256 1 TRYD_DROER

P54625 drosophila
P35004 drosophila
P54626 drosophila
P42276 drosophila
P42277 drosophila
P29786 aedes aegypti
P51588 sarcophaga
P35044 lucilia cup
P35587 hypoderma l
P35388 hypoderma l
P54624 drosophila
P04814 drosophila

ALIGNMENTS

RESULT 1

THRB_RAT
ID THRB_RAT STANDARD; PRT; 617 AA.
AC P18292;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Prothrombin precursor (EC 3.4.21.5).
GN F2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=90332426; PubMed=2377469;
RA Dihaich M., Monard D.;
RT "CDNA sequence of rat prothrombin";
RL Nucleic Acids Res. 18:4251-4251(1990).
RN [2]
RP SEQUENCE OF 383-617 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., Macgillivray R.T.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT amplification and sequence analysis of the B chain of thrombin from
RT nine different species";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC fibrinogen to fibrin and releases fibrinopeptide A and B.
CC -1- RESULT FROM THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOAMAL
CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
CC OF PROTHROMBIN TO THROMBIN.
CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC THROMBIN.
CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC BY FACTOR XA.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; X52835; CAA37017.1; -;
EMBL; M81397; AAN42240.1; -;
PIR; S10511; S10511.
DR HSPG; P00734; LUVS.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002383; GLA blood.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00594; gla; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00001; GLABLOOD.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00069; GLA; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW Vitamin K; Serine; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW Hydrolase; Zymogen; Gamma-carboxyglutamic acid; Kringle; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 43
FT CHAIN 44 617 PROTHROMBIN.
FT PEPTIDE 44 200
FT ACT_SITE 201 323 ACTIVATION PEPTIDE (FRAGMENT 1).
FT CHAIN 324 359 ACTIVATION PEPTIDE (FRAGMENT 2).
FT CHAIN 360 617 THROMBIN LIGHT CHAIN (A).
FT DOMAIN 109 187 THROMBIN HEAVY CHAIN (B).
FT DOMAIN 215 292 KRINGLE 1.
FT DOMAIN 360 617 KRINGLE 2.
FT SITE 200 201 SERINE PROTEASE.
FT SITE 323 324 CLEAVAGE (BY THROMBIN).
FT SITE 359 360 CLEAVAGE (BY FACTOR XA).
FT ACT_SITE 402 402 CLEAVAGE (BY FACTOR XA).
FT ACT_SITE 458 458 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 564 564 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT MOD_RES 50 50 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 51 51 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 58 58 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 63 63 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 64 64 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 76 76 GAMMA-CARBOXYGLUTAMIC ACID.
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 552 552 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 61 66 BY SIMILARITY.
FT DISULFID 91 104 BY SIMILARITY.
FT DISULFID 109 187 BY SIMILARITY.
FT DISULFID 130 170 BY SIMILARITY.
FT DISULFID 158 182 BY SIMILARITY.
FT DISULFID 215 292 BY SIMILARITY.
FT DISULFID 236 276 BY SIMILARITY.

FT DISULFID 264 287 BY SIMILARITY.
FT DISULFID 332 478 INTERCHAIN (BY SIMILARITY).
FT DISULFID 387 403 BY SIMILARITY.
FT DISULFID 532 546 BY SIMILARITY.
FT DISULFID 560 590 BY SIMILARITY.
SQ SEQUENCE 617 AA; 70411 MW; AD27DIB71445DBID CRC64;
Query Match 100.0%; Score 59; DB 1; Length 617;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEGSGGPFV 10
Db 560 CEGSGGPFV 569
RESULT 2
THRB_MOUSE
ID THRB_MOUSE STANDARD; PRT; 618 AA.
AC P19221;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Prothrombin precursor (EC 3.4.21.5).
DE F2 OR CF2.
GN F2 OR CF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=91025551; PubMed=2222810;
RA Friesner Degen S.J., Schaffer L.A., Jamison C.S., Grant S.G.,
RA Fitzgibbon J.J., Pai J.-A., Chapman V.M., Elliott R.W.;
RT "Characterization of the cDNA coding for mouse prothrombin and
RT localization of the gene on mouse chromosome 2.";
RL DNA Cell Biol. 9:487-498(1990).
RN [2]
RN SEQUENCE OF 384-618 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., Macgillivray R.T.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC fibrinogen to fibrin and releases fibrinopeptide A and B.
CC -1- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
CC OF PROTHROMBIN TO THROMBIN.
CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC THROMBIN.
CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC BY FACTOR XA.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X52308; CAA36548.1; .
EMBL; M81394; AAA04035.1; .
PIR; A35827; A35827.
HSSP; P00734; 1B7X.
MEROPS; S01.217; .
MGD; MGI:88380; F2.
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR002383; GLA_blood.
InterPro; IPR000001; Kringle.
InterPro; IPR003966; Prothrombin.
InterPro; IPR001254; Ser protease_Try.
InterPro; IPR000294; vitK_dep_GLA.
Pfam; PF00051; kringle; 2.
Pfam; PF00089; trypsin; 1.
Pfam; PF00594; gla; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00001; GLABLOOD.
PRINTS; PR00018; KRINGLE.
PRINTS; PR01505; PROTHROMBIN.
ProDom; PD000395; Kringle; 2.
SMART; SM00069; GLA; 1.
SMART; SM00130; KR; 2.
SMART; SM00020; TRYP_SPE; 1.
PROSITE; PS00011; GIL CARBOXYLATION; 1.
PROSITE; PS00021; KRINGLE_1; 2.
PROSITE; PS00070; KRINGLE_2; 2.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
Hydrolase; Serine protease; Kringle; Signal.
SIGNAL 1 24
PROPEP 25 43
CHAIN 44 618
PEPTIDE 44 200
ACTIVATION PEPTIDE (FRAGMENT 1).
ACTIVATION PEPTIDE (FRAGMENT 2).
THROMBIN LIGHT CHAIN (A).
THROMBIN HEAVY CHAIN (B).
KRINGLE 1.
KRINGLE 2.
SERINE PROTEASE.
CLEAVAGE (BY THROMBIN).
CLEAVAGE (BY FACTOR XA).
CLEAVAGE (BY FACTOR XA).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
GAMMA-CARBOXYGLUTAMIC ACID.
GAMMA-CARBOXYGLUTAMIC ACID.
GAMMA-CARBOXYGLUTAMIC ACID.
GAMMA-CARBOXYGLUTAMIC ACID.
GAMMA-CARBOXYGLUTAMIC ACID.
GAMMA-CARBOXYGLUTAMIC ACID.
GAMMA-CARBOXYGLUTAMIC ACID.
GAMMA-CARBOXYGLUTAMIC ACID.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
INTERCHAIN (BY SIMILARITY).

FT DISULFID 388 404 BY SIMILARITY.
FT DISULFID 533 547 BY SIMILARITY.
FT DISULFID 561 591 BY SIMILARITY.
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .).
FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .).
FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .).
SQ SEQUENCE 618 AA; 70268 MW; B89F719A9FD601E0 CRC64;
Query Match 100.0%; Score 59; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEGDSGGPFV 10
DB 561 CEGDSGGPFV 570
RESULT 3
THRB_HUMAN
ID THRB_HUMAN STANDARD; PRT; 622 AA.
AC P00734;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II).
GN F2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88077877; PubMed=2825773;
RA Degen S.J.F., Davie E.W.;
RT "Nucleotide sequence of the gene for human prothrombin.";
RL Biochemistry 26:6165-6177(1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT MET-165.
RA Rieder M.J., Armet T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
RA Ozuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 8-622 FROM N.A.
RX MEDLINE=83231469; PubMed=6305407;
RA Degen S.J.F., McGilivray R.T.A., Davie E.W.;
RT "Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin.";
RL Biochemistry 22:2087-2097(1983).
RN [4]
RP SEQUENCE OF 44-314.
RX MEDLINE=77193964; PubMed=266717;
RA Walz D.A., Hewett-Emmett D., Seegers W.H.;
RT "Amino acid sequence of human prothrombin fragments 1 and 2.";
RL Proc. Natl. Acad. Sci. U.S.A. 74:1969-1972(1977).
RN [5]
RP SEQUENCE OF 315-622.
RX MEDLINE=77207112; PubMed=873923;
RA Rutkowski R.J., Elion J., Downing M.R., Mann K.G.;
RT "Primary structure of human prothrombin 2 and alpha-thrombin.";
RL J. Biol. Chem. 252:4942-4957(1977).
RN [6]
RP PROCESSING.
RX MEDLINE=87008532; PubMed=3759958;
RA Rabiet M.J., Blashill A., Furie B., Furie B.C.;
RT "Prothrombin fragment 1 X 2 X 3, a major product of prothrombin activation in human plasma.";
RL J. Biol. Chem. 261:13210-13215(1986).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=90059942; PubMed=2583108;
RA Bode W., Mayr I., Baumann U., Huber R., Stone S.R., Hofsteenge J.;
RT "The refined 1.9 A crystal structure of human alpha-thrombin:

RT interaction with D-Phe-Pro-Arg chloromethylketone and significance of
RT the Tyr-Pro-Pro-Trp insertion segment.";
RL EMBO J. 8:3467-3475(1989).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE-90327074; PubMed-2374926;
RA Rydel T.J., Ravichandran K.G., Tulinsky A., Bode W., Huber R.,
RA Roltsch C., Fenton J.W. II;
RT "The structure of a complex of recombinant hirudin and human alpha-
RT thrombin.";
RL Science 249:277-280(1990).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE-94330942; PubMed-8071320;
RA Rydel T.J., Yin M., Padmanabhan K.P., Blankenship D.T., Cardin A.D.,
RA Correa P.E., Fenton J.W. II, Tulinsky A.;
RT "Crystallographic structure of human gamma-thrombin.";
RL J. Biol. Chem. 269:22000-22006(1994).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE-97357286; PubMed-9214615;
RA van de Locht A., Bode W., Huber R., le Bonniec B.F., Stone S.R.,
RA Esmen C.T., Stubbs M.T.;
RT "The thrombin E192Q-BPTI complex reveals gross structural
RT rearrangements: implications for the interaction with antithrombin
RT and thrombomodulin.";
RL EMBO J. 16:2977-2984(1997).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.
RX MEDLINE-99162521; PubMed-10051558;
RA Guinto E.R., Caccia S., Rose T., Fuetterer K., Wakaman G., di Cera E.;
RT "Unexpected crucial role of residue 225 in serine proteases.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857(1999).
RN [12]
RP VARIANT BARCELONA.
RX MEDLINE-87033739; PubMed-3771562;
RA Rabiet M.-J., Furie B.C., Furie B.;
RT "Molecular defect of prothrombin Barcelona. Substitution of cysteine
RT for arginine at residue 273.";
RL J. Biol. Chem. 261:15045-15048(1986).
RN [13]
RP VARIANT FRANKFURT.
RX MEDLINE-95313001; PubMed-7792730;
RA Degen S.J.F., McDowell S.A., Sparks L.M., Scharrer I.;
RT "Prothrombin Frankfurt: a dysfunctional prothrombin characterized by
RT substitution of Glu-466 by Ala.";
RL Thromb. Haemost. 73:203-209(1995).
RN [14]
RP VARIANTS HIMI-1 AND HIMI-2.
RX MEDLINE-93043342; PubMed-1421398;
RA Morishita E., Saito M., Kumabashiri I., Asakura H., Matsuda T.,
RA Yanaguchi K.;
RT "Prothrombin Himi: a compound heterozygote for two dysfunctional
RT prothrombin molecules (Met-337-->Thr and Arg-388-->His).";
RL Blood 80:2275-2280(1992).
RN [15]
RP VARIANT PADUA-1.
RX MEDLINE-9516998; PubMed-7865694;
RA James H.L., Kim D.J., Zheng D.-Q., Girolami A.;
RT "Prothrombin Padua I: incomplete activation due to an amino acid
RT substitution at a factor Xa cleavage site.";
RL Blood Coagul. Fibrinolysis 5:841-844(1994).
RN [16]
RP VARIANT QUICK-1.
RX MEDLINE-89207504; PubMed-3342619;
RA Henriksen R.A., Mann K.G.;
RT "Identification of the primary structural defect in the dysfibrinogen
RT thrombin Quick I: substitution of cysteine for arginine-382.";
RL Biochemistry 27:9160-9165(1988).
RN [17]
RP VARIANT QUICK-2.
RX MEDLINE-89247398; PubMed-2719946;
RA Henriksen R.A., Mann K.G.;

RT "Substitution of valine for glycine-558 in the congenital dysfibrinogen
RT thrombin Quick II alters primary substrate specificity.";
RL Biochemistry 28:2078-2082(1989).
RN [18]
RP VARIANT SALAKTA.
RX MEDLINE-92378975; PubMed-1354985;
RA Miyata T., Aruga R., Uneyama H., Bezeaud A., Guillin M.-C.,
RA Iwanaga S.;
RT "Prothrombin Salakta: substitution of glutamic acid-466 by alanine
RT reduces the fibrinogen clotting activity and the esterase activity.";
RL Biochemistry 31:7457-7462(1992).
RN [19]
RP VARIANT TOKUSHIMA.
RX MEDLINE-87185407; PubMed-3567158;
RA Miyata T., Morita T., Inomoto T., Kawauchi S., Shirakami A.,
RA Iwanaga S.;
RT "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan
RT that impairs the fibrinogen clotting activity of derived thrombin
RT Tokushima.";
RL Biochemistry 26:1117-1122(1987).
RN [20]
RP VARIANT TOKUSHIMA.
RX MEDLINE-87101511; PubMed-3801671;
RA Inomoto T., Shirakami A., Kawauchi S., Shigeakiyo T., Saito S.,
RA Miyoshi K., Morita T., Iwanaga S.;
RT "Prothrombin Tokushima: characterization of dysfunctional thrombin
RT derived from a variant of human prothrombin.";
RL Blood 69:565-569(1987).
RN [21]
RP VARIANT TOKUSHIMA.
RX MEDLINE-92256895; PubMed-1349838;
RA Iwahana H., Yoshimoto K., Shigeakiyo T., Shirakami A., Saito S.,
RA Itakura M.;
RT "Detection of a single base substitution of the gene for prothrombin
RT Tokushima. The application of PCR-SSCP for the genetic and molecular
RT analysis of dysprothrombinemia.";
RL Int. J. Hematol. 55:93-100(1992).
RN [22]
RP VARIANT TYPE-3.
RX MEDLINE-83204687; PubMed-6405779;
RA Board P.G., Shaw D.C.;
RT "Determination of the amino acid substitution in human prothrombin
RT type 3 (157 Glu leads to Lys) and the localization of a third
RT thrombin cleavage site.";
RL Br. J. Haematol. 54:245-254(1983).
CC -|- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC fibrinogen to fibrin and releases fibrinopeptide A and B.
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
CC -|- PM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
CC OF PROTHROMBIN TO THROMBIN.
CC -|- DISEASE: DEFECTS IN F2 ARE THE CAUSE OF VARIOUS FORMS OF
CC DYSPROTHROMBINEMIA.
CC -|- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC THROMBIN.
CC -|- MISCELLANEOUS: IT IS NOT KNOWN WHETHER 1 OR 2 SMALLER ACTIVATION
CC PEPTIDES, WITH ADDITIONAL CLEAVAGE AFTER 314-ARG, ARE RELEASED IN
CC NATURAL BLOOD CLOTTING.
CC -|- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC BY FACTOR XA.

CC -1- MISCELLANEOUS: THE CLEAVAGE AFTER R-198, OBSERVED IN VITRO, DOES

Query Match 100.0%; Score 59; DB 1; Length 622;
 Best Local Similarity 100.0%; Pred. No. 0.0062;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
 |||||
 DB 564 CEGDSGGPFV 573

RESULT 4
 THRB_BOVIN STANDARD; PRT; 625 AA.
 AC P00735;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Prothrombin precursor (EC 3.4.21.5).
 GN F2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8245190; PubMed=3379642;
 RA Irwin D.M., Robertson K.A., Macgillivray R.T.A.;
 RT "Structure and evolution of the bovine prothrombin gene.";
 RL J. Mol. Biol. 200;31-45(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84203525; PubMed=6326805;
 RA McGillivray R.T.A., Davie E.W.;
 RT "Characterization of bovine prothrombin mRNA and its translation product.";
 RL Biochemistry 23:1626-1634(1984).
 RN [3]
 RP SEQUENCE OF 44-625, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
 RA Magnusson S., Sottrup-Jensen L., Petersen T.E., Claeys H.;
 RL (In) Hemker H.C., Veitkamp J.J. (eds.);
 RL Boerhaave symposium on prothrombin and related coagulation factors,
 RL pp.25-46, Leiden University Press, Leiden (1975).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RX MEDLINE=86296631; PubMed=3741841;
 RA Park C.H., Tulinsky A.;
 RT "Three-dimensional structure of the kringle sequence: structure of prothrombin fragment 1.";
 RL Biochemistry 25:3977-3982(1986).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RX MEDLINE=91311686; PubMed=1856869;
 RA Seshadri T.-P., Tulinsky A., Skrzypczak-Jankun E., Park C.H.;
 RT "Structure of bovine prothrombin fragment 1 refined at 2.25-A resolution.";
 RL J. Mol. Biol. 220:481-494(1991).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RX MEDLINE=92190185; PubMed=1547238;
 RA Soriano-Garcia M., Padmanabhan K., de Vos A.M., Tulinsky A.;
 RT "The Ca2+ ion and membrane binding structure of the Gla domain of Ca-prothrombin fragment 1.";
 RL Biochemistry 31:2554-2566(1992).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=92218459; PubMed=1560020;
 RA Martin P.D., Robertson W., Turk D., Huber R., Bode W., Edwards B.F.P.;
 RT "The structure of residues 7-16 of the A alpha-chain of human fibrinogen bound to bovine thrombin at 2.3-A resolution.";
 RL J. Biol. Chem. 267:7911-7920(1992).
 RN [8]

RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=92289319; PubMed=1518046;
 RA Brandstetter H., Turk D., Hoeffken H.W., Grosse D., Stuerzebecher J.,
 RA Martin P.D., Edwards B.F.P., Bode W.;
 RT "Refined 2.3 A X-ray crystal structure of bovine thrombin complexed with the benzamidine and arginine-based thrombin inhibitors NAPAP, 4-TAPAP and MQPA. A starting point for improving antithrombotics.";
 RL J. Mol. Biol. 226:1085-1089(1992).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.
 RX MEDLINE=97102783; PubMed=8947023;
 RA van de Locht A., Stubbs M.T., Bode W., Friedrich T., Bollschweiler C.,
 RA Hoffken W., Huber R.;
 RT "The ornithodorin-thrombin crystal structure, a key to the TAP enigma?";
 RL EMBO J. 15:6011-6017(1996).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH TRIABIN.
 RX MEDLINE=98004486; PubMed=9342325;
 RA Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,
 RA Huber R., Bode W.;
 RT "Structure of the thrombin complex with triabin, a lipocalin-like exosite-binding inhibitor derived from a triatomine bug.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11845-11850(1997).
 RN [11]
 RP GENE STRUCTURE.
 RX MEDLINE=86077733; PubMed=3000440;
 RA Irwin D.M., Ahern K.G., Pearson G.D., McGillivray R.T.A.;
 RT "Characterization of the bovine prothrombin gene.";
 RL Biochemistry 24:6854-6861(1985).
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII, AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates fibrinogen to fibrin and releases fibrinopeptide A and B.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- PM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS, RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION OF PROTHROMBIN TO THROMBIN.
 CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN & FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT & HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF THROMBIN.
 CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION BY FACTOR XA.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC -1- DATABASE: NAME-Prozyme technical fact sheet;
 CC WWW="http://www.prozyme.com/technical/thrombindata.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V00135; CAA23451.1; -
 DR EMBL; J00041; AAA30781.1; -
 DR PIR; A00915; TBBO.
 DR PIR; S02537; S02537.
 DR PDB; 1BBR; 31-JAN-94.
 DR PDB; 1ETR; 31-JAN-94.


```

QY 1 CEGDSGGPFV 10
Db 348 CQDSDGGPFV 357

RESULT 6
HEPS_HUMAN STANDARD; PRT; 417 AA.
ID HEPS_HUMAN STANDARD; PRT; 417 AA.
AC P05981;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine protease hepsin (EC 3.4.21.-) (transmembrane protease, serine
DE 1).
DE HPN OR TMPPRS1.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC MEDLINE=91358502; PubMed=2835076;
RX Leytus S.P., Loeb K.R., Hagen F.S., Kurachi K., Davie E.W.;
RT "A novel trypsin-like serine protease (hepsin) with a putative
RL transmembrane domain expressed by human liver and hepatoma cells.";
RL Biochemistry 27:1067-1074(1988).
RN [2]
RN CHARACTERIZATION.
RP MEDLINE=91358502; PubMed=1885621;
RX Tsuji A., Torres-Rosado A., Arai T., le Beau M.M., Lemons R.S.,
RA Chou S.H., Kurachi K.;
RT "Hepsin, a cell membrane-associated protease. Characterization,
RL tissue distribution, and gene localization.";
RL J. Biol. Chem. 266:16948-16953(1991).
RN [3]
RN CHARACTERIZATION.
RP MEDLINE=93348237; PubMed=8346233;
RX Torres-Rosado A., O'Shea K.S., Tsuji A., Chou S.H., Kurachi K.;
RA "Hepsin, a putative cell-surface serine protease, is required for
RL mammalian cell growth.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7181-7187(1993).
CC -1- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CELL GROWTH AND MAINTENANCE
CC OF CELL MORPHOLOGY.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: PRESENT IN MOST TISSUES, WITH THE HIGHEST
CC LEVEL IN LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M18930; AAA36013.1; -
CC EMBL; X07732; CAA30558.1; -
CC EMBL; X07002; CAA30058.1; -
CC PIR; S00845; S00845.
CC HSP; P00763; IDPO.
CC MEROPS; S01.224; -
CC Genew; HGNC:5155; HPN.
CC MIM; 142440; -
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC InterPro; IPR001190; Srrc_receptor.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00202; SR; 1.
CC SMART; SM00020; Tryp_SPC; 1.

QY 1 CEGDSGGPFV 10
Db 349 CQDSDGGPFV 358

RESULT 7
PRTC_CANFA STANDARD; PRT; 157 AA.
ID PRTC_CANFA STANDARD; PRT; 157 AA.
AC Q28278;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)
DE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).
CN PROC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=94318474; PubMed=8043441;
RX Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RA "A comparative study of partial primary structures of the catalytic
RL region of mammalian protein C.";
RN Br. J. Haematol. 86:590-600(1994).
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC and VIIIa.
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D43751; BAA07808.1; -
CC HSP; P04070; IPCU.
CC MEROPS; S01.218; -
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00089; trypsin; 1.
CC SMART; SM00020; Tryp_SPC; 1.

Query Match 94.9%; Score 56; DB 1; Length 417;
Best Local Similarity 90.0%; Pred. No. 0.014;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CEGDSGGPFV 10
Db 349 CQDSDGGPFV 358

```

DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
 FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
 FT ACT_SITE 125 125 CHARGE RELAY SYSTEM.
 FT DISULFID 96 110 BY SIMILARITY.
 FT DISULFID 121 149 BY SIMILARITY.
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 157 157
 SQ SEQUENCE 157 AA; 17262 MW; B8B1BACF49220DFB CRC64;
 Query Match 89.8%; Score 53; DB 1; Length 157;
 Best Local Similarity 90.0%; Pred. No. 0.018;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
 |||||
 Db 121 CEGDSGGPMV 130

RESULT 8
 PRTC_CAPHI
 ID PRTC_CAPHI STANDARD; PRT; 157 AA.
 AC Q28315;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)
 DE Anticoagulant protein C (Blood coagulation factor XIV) (Fragment).
 GN PROC.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94318474; PubMed=8043441;
 RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
 RT "A comparative study of partial primary structures of the catalytic
 region of mammalian protein C.";
 RL Br. J. Haematol. 86:590-600(1994).
 CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
 REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
 IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
 CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
 and VIIIA.
 CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D43752; BAA07809.1; -
 CC HSP; P04070; IPCU.
 CC MEROPS; S01.218; -
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; TRYPSIN_DOM; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
 FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
 FT ACT_SITE 125 125 CHARGE RELAY SYSTEM.
 FT DISULFID 96 110 BY SIMILARITY.
 FT DISULFID 121 149 BY SIMILARITY.
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 157 157
 SQ SEQUENCE 157 AA; 17142 MW; FBDC5BE2ECA74BB4 CRC64;

FT ACT_SITE 125 125 CHARGE RELAY SYSTEM.
 FT DISULFID 96 110 BY SIMILARITY.
 FT DISULFID 121 149 BY SIMILARITY.
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 157 157
 SQ SEQUENCE 157 AA; 17251 MW; B89790F9954B610A CRC64;
 Query Match 89.8%; Score 53; DB 1; Length 157;
 Best Local Similarity 90.0%; Pred. No. 0.018;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
 |||||
 Db 121 CEGDSGGPMV 130

RESULT 9
 PRTC_FELCA
 ID PRTC_FELCA STANDARD; PRT; 157 AA.
 AC Q28412;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)
 DE Anticoagulant protein C (Blood coagulation factor XIV) (Fragment).
 GN PROC.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94318474; PubMed=8043441;
 RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
 RT "A comparative study of partial primary structures of the catalytic
 region of mammalian protein C.";
 RL Br. J. Haematol. 86:590-600(1994).
 CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
 REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
 IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
 CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
 and VIIIA.
 CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D43750; BAA07807.1; -
 CC HSP; P04070; IPCU.
 CC MEROPS; S01.218; -
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; TRYPSIN_DOM; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
 FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
 FT ACT_SITE 125 125 CHARGE RELAY SYSTEM.
 FT DISULFID 96 110 BY SIMILARITY.
 FT DISULFID 121 149 BY SIMILARITY.
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 157 157
 SQ SEQUENCE 157 AA; 17142 MW; FBDC5BE2ECA74BB4 CRC64;

Query Match 89.8%; Score 53; DB 1; Length 157;
 Best Local Similarity 90.0%; Pred. No. 0.018;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
 |||||

DB 121 CEGDSGGPMV 130

RESULT 10

PTC_HORSE STANDARD; PRT; 157 AA.
 AC Q28380;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)
 DE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).
 GN PROC.
 OS Equus caballus (Horse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OC NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94318474; PubMed=8043441;
 RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
 RT "A comparative study of partial primary structures of the catalytic
 region of mammalian protein C";
 RL Br. J. Haematol. 86:590-600(1994).
 CC - FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
 REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
 IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
 CC - CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
 and VIIIA.
 CC - TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
 CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

DR EMBL; D43753; BAA07810.1; -
 DR HSP; P04070; LPCU.
 DR MEROPS; S01.218; -
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00240; TRYP_SIN_DOM; 1.
 DR PROSITE; PS00134; TRYP_SIN_HIS; PARTIAL.
 DR PROSITE; PS00135; TRYP_SIN_SER; 1.
 DR Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
 KW NON_TER 1
 FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
 FT ACT_SITE 125 125 CHARGE RELAY SYSTEM.
 FT DISULFID 96 110 BY SIMILARITY.
 FT DISULFID 121 149 BY SIMILARITY.
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 157 157
 SQ SEQUENCE 157 AA; 17200 MW; BFA6EA045C3C580 CRC64;

Query Match 89.8%; Score 53; DB 1; Length 157;
 Best Local Similarity 90.0%; Pred. No. 0.018;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
 |||||

DB 121 CEGDSGGPMV 130

RESULT 11

PTC_MACMU STANDARD; PRT; 161 AA.
 AC Q28506;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)
 DE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).
 GN PROC.
 OS Macaca mulatta (Rhesus macaque).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OC NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94318474; PubMed=8043441;
 RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
 RT "A comparative study of partial primary structures of the catalytic
 region of mammalian protein C";
 RL Br. J. Haematol. 86:590-600(1994).
 CC - FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
 REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
 IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
 CC - CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
 and VIIIA.
 CC - TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
 CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

DR EMBL; D43754; BAA07811.1; -
 DR HSP; P04070; LPCU.
 DR MEROPS; S01.218; -
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00240; TRYP_SIN_DOM; 1.
 DR PROSITE; PS00134; TRYP_SIN_HIS; PARTIAL.
 DR PROSITE; PS00135; TRYP_SIN_SER; 1.
 DR Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
 KW NON_TER 1
 FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
 FT ACT_SITE 129 129 CHARGE RELAY SYSTEM.
 FT DISULFID 100 114 BY SIMILARITY.
 FT DISULFID 125 153 BY SIMILARITY.
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 161 161
 SQ SEQUENCE 161 AA; 17770 MW; 27D78F185B2FCC69 CRC64;

Query Match 89.8%; Score 53; DB 1; Length 161;
 Best Local Similarity 90.0%; Pred. No. 0.018;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEGDSGGPFV 10
 |||||

DB 125 CEGDSGGPMV 134

RESULT 12

KLKF_HUMAN STANDARD; PRT; 256 AA.
 ID KLKF_HUMAN

AC Q9H2R5: Q9H2R6; Q9H2R4; Q9H2R3; Q9HBG9; Q15358;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kallikrein 15 precursor (EC 3.4.21.-) (ACO protease).
 GN KKL15.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_Taxid=9606;
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RX PubMed=11010966;
 RA Yousef G.M., Scorilas A., Jung K., Ashworth L.K., Diamandis E.P.;
 RT "Molecular cloning of the human kallikrein 15 gene (KLK15). Up-
 regulation in prostate cancer.";
 RL J. Biol. Chem. 276:53-61(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20510030; PubMed=11054574;
 RA Can L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
 RA Moss P., Paepker B., Wang K.;
 RT "Sequencing and expression analysis of the serine protease gene
 cluster located in chromosome 19q13 region.";
 RL Gene 257:119-130(2000).
 RN [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94289486; PubMed=8018728;
 RA Dhanich M.E., Spiess M.;
 RT "A novel serine proteinase-like sequence from human brain.";
 RL Biochim. Biophys. Acta 1218:225-228(1994).
 CC -1- FUNCTION: Protease whose physiological substrate is not yet known.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Highest expression in the thyroid gland. Also
 CC expressed in the prostate, salivary, and adrenal glands and in the
 CC colon testis and kidney.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF242195; AAG09469.1;
 CC EMBL; AF242195; AAG09470.1;
 CC EMBL; AF242195; AAG09471.1;
 CC EMBL; AF242195; AAG09472.1;
 CC EMBL; AF243527; AAG33354.1;
 CC EMBL; X75363; CAA53145.1; ALT_SEQ.
 CC HSP; P00763; IDPO.
 CC MEROPS: S01.081.
 CC InterPro: IPR001314; Chymotrypsin.
 CC InterPro: IPR001254; Ser_protease_Try.
 CC Pfam: PF00089; trypsin.1.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM00020; TRYPSIN_DOM; 1.
 CC PROSITE: PS0240; TRYPSIN_DOM; 1.
 CC PROSITE: PS00134; TRYPSIN_HIS; 1.
 CC PROSITE: PS00135; TRYPSIN_SER; FALSE_NEG.
 CC HydroLase: Serine protease; Glycoprotein; Signal; Zymogen;
 CC Alternative splicing.
 CC SIGNAL 1 16 POTENTIAL.
 CC PROPEP 17 21
 CC CHAIN 22 256
 CC ACT_SITE 62 62
 CC ACT_SITE 106 106
 CC ACT_SITE 209 209

FT CARBOHYD 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 122 MISSING (IN ISOFORM 4).
 FT VARSPLIC 122 MISSING (IN ISOFORM 2).
 FT VARSPLIC 161 V -> G (IN ISOFORM 3).
 FT VARSPLIC 162 MISSING (IN ISOFORM 3).
 FT CONFLICT 147 SHNEPGTAGSPRSQ -> PLSSP (IN REF. 2).
 SQ SEQUENCE 256 AA; 28087 MW; B5EBFD6022786B5 CRC64;
 Query Match 89.8%; Score 53; DB 1; Length 256;
 Best Local Similarity 90.0%; Pred. No. 0.029;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CEGDSGGPFV 10
 DB 205 CEGDSGGPLY 214
 RESULT 13
 KLLD_HUMAN STANDARD; PRT; 277 AA.
 AC Q9UKR3; Q9Y433;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Kallikrein 13 precursor (EC 3.4.21.-) (Kallikrein-like protein 4).
 DE (KLK-L4).
 GN KLK13 OR KLK14.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20229789; PubMed=10766816;
 RA Yousef G.M., Chang A., Diamandis E.P.;
 RT "Identification and characterization of KLK-L4, a new kallikrein-like
 RT gene that appears to be down-regulated in breast cancer tissues.";
 RL J. Biol. Chem. 275:11891-11898(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Garnes J.,
 RA Danganan L., Eriar A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
 RA Olsen A.S., Carrano A.V.;
 RT "Sequence analysis of chromosome 19q13.4.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-180 FROM N.A.
 RC TISSUE=Uterus;
 RA Ansoorge W., Wirkner U., Meves H.-W., Gassenhuber J., Wiemann S.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN PROSTATE, BREAST, TESTIS AND
 CC SALIVARY GLAND.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF135024; AAD26425.2;
 CC EMBL; AC011473; AAG23259.1;
 CC EMBL; AL050220; CAB43320.1; ALT_INIT.
 CC HSP; P00763; IDPO.
 DR

DR MEROPS; S01.306; -
 DR Gnew; HGNC:6361; KLR13.
 DR MIM; 605505; -
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYPSIN_DOM; 1.
 DR PROSITE; PS00240; TRYPSIN_HIS; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease; Glycoprotein; Signal.
 KW SIGNAL 1 16
 FT CHAIN 17 277
 FT ACT_SITE 76 76
 FT ACT_SITE 124 124
 FT ACT_SITE 218 218
 FT ACT_SITE 218 218
 FT DISULFID 42 178
 FT DISULFID 61 77
 FT DISULFID 157 224
 FT DISULFID 189 203
 FT DISULFID 214 239
 FT CARBOHYD 30 30
 FT CARBOHYD 225 225
 FT CONFLICT 170 180
 FT SEQUENCE 277 AA; 30570 MW; BABA9E8DCFB5D542 CRC64;
 Query Match 89.8%; Score 53; DB 1; Length 277;
 Best Local Similarity 90.0%; Pred. No. 0.031;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CEGDGGGPV 10
 DB 214 CEGDGGPLV 223
 RESULT 14
 PRTC_BOVIN STANDARD; PRT; 456 AA.
 AC P00745;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vitamin-K dependent protein C precursor (SC 3.4.21.69)
 DE (Autoproteolysis IIA) (Anticoagulant protein C) (Blood coagulation
 factor XIV) (Fragment).
 GN PROC.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-85014826; PubMed-6091100;
 RA Long G.L., Balagaje R.M., McGillivray R.T.A.;
 RT "Cloning and sequencing of liver cDNA coding for bovine protein C.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5653-5656(1984).
 RN [2]
 RP SEQUENCE OF 40-194.
 RX MEDLINE-83007325; PubMed-6896876;
 RA Fernlund P., Stenflo J.;
 RT "Amino acid sequence of the light chain of bovine protein C.";
 RL J. Biol. Chem. 257:12170-12179(1982).
 RN [3]
 RP REVISION TO 110.
 RX MEDLINE-83169769; PubMed-6572939;
 RA Drakenberg T., Fernlund P., Roepstorff P., Stenflo J.;
 RT "Beta-hydroxyaspartic acid in vitamin K-dependent protein C.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:1802-1806(1983).
 RN [4]
 RP SEQUENCE OF 197-456.
 RX MEDLINE-83007326; PubMed-6896877;

RA Stenflo J., Fernlund P.;
 RT "Amino acid sequence of the heavy chain of bovine protein C.";
 RL J. Biol. Chem. 257:12180-12190(1982).
 RN [5]
 RP PROCESSING AND CALCIUM-BINDING DATA.
 RX MEDLINE-83213513; PubMed-6304092;
 RA Esmon N.L., Debault L.E., Esmon C.T.;
 RT Proteolytic formation and properties of gamma-carboxyglutamic acid-
 domainless protein C.;
 RL J. Biol. Chem. 258:5548-5553(1983).
 RN [6]
 RP PROCESSING AND CALCIUM-BINDING DATA.
 RX MEDLINE-83213514; PubMed-6406503;
 RA Johnson A.E., Esmon N.L., Laue T.M., Esmon C.T.;
 RT "Structural changes required for activation of protein C are induced
 by Ca2+ binding to a high affinity site that does not contain gamma-
 carboxyglutamic acid.";
 RL J. Biol. Chem. 258:5554-5560(1983).
 CC 1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
 REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
 IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
 CC 1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
 and VIIIA.
 CC 1- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
 INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
 BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
 TETRAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS
 REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
 STRONGLY PROMOTED BY THROMBOMODULIN.
 CC 1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
 CC 1- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME
 GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
 CC 1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
 ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
 SITE IS NECESSARY FOR THE RECOGNITION OF THE
 THROMBIN-THROMBOMODULIN COMPLEX.
 CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC 1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K02435; AAA30685.1; -
 DR PIR; A00928; KXBO.
 DR HSP; P04070; IPCU.
 DR MEROPS; S01.218; -
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR InterPro; IPR000294; VitK_dep_GLA.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR Pfam; PF00594; gla; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00069; GLA; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_Ca; 1.
 DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Blood coagulation; Glycoprotein; Serine protease;
 KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
 EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.

FT NON_TER 1
FT SIGNAL <1
FT PROPEP 30
FT CHAIN 194
FT CHAIN 456
FT PEPTIDE 197
FT PEPTIDE 210
FT DOMAIN 94
FT DOMAIN 129
FT DOMAIN 133
FT DOMAIN 173
FT DOMAIN 211
FT MOD_RES 45
FT MOD_RES 46
FT MOD_RES 53
FT MOD_RES 55
FT MOD_RES 58
FT MOD_RES 59
FT MOD_RES 62
FT MOD_RES 64
FT MOD_RES 65
FT MOD_RES 68
FT MOD_RES 74
FT MOD_RES 110
FT ACT_SITE 232
FT ACT_SITE 298
FT ACT_SITE 397
FT ACT_SITE 397
FT DISULFID 56
FT DISULFID 89
FT DISULFID 98
FT DISULFID 102
FT DISULFID 119
FT DISULFID 137
FT DISULFID 148
FT DISULFID 157
FT DISULFID 172
FT DISULFID 180
FT DISULFID 237
FT DISULFID 368
FT DISULFID 393
FT CARBOHYD 136
FT CARBOHYD 289
FT CARBOHYD 350
FT CARBOHYD 366
FT VARIANT 82
FT CONFLICT 455
SQ SEQUENCE 456 AA: 51407 MW: CAAAF6833F894C209 CRC64;
Query Match 89.8%; Score 53; DB 1; Length 456;
Best Local Similarity 90.0%; Pred. NO. 0.052;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CREGDSGPFV 10
DB 393 CREGDSGPMV 402
RESULT 15
ID PRTC_PIG STANDARD; PRT; 459 AA.
AC Q9GLP2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)
DE (Autoproteolytic protein C precursor (EC 3.4.21.69)
DE factor XIV).
GN PROC.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21121490; PubMed=11229814;
RA Grimm D.R., Colter M.B., Braunschweig M., Alexander L.J., Neame P.J.,

RA Kim H.K.W.;
RT Porcine factor V: cDNA cloning, gene mapping, three-dimensional
RT protein modeling of membrane binding sites and comparative anatomy of
RT domains.";
RL Cell. Mol. Life Sci. 58:148-159(2001).
CC -!- FUNCTION: Protein C is a vitamin K-dependent serine protease that
CC regulates blood coagulation by inactivating factors Va and VIIa
CC in the presence of calcium ions and phospholipids.
CC -!- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC and VIIa.
CC -!- SUBUNIT: Synthesized as a single chain precursor, which is cleaved
CC into a light chain and a heavy chain held together by a disulfide
CC bond. The enzyme is then activated by thrombin, which cleaves a
CC tetradecapeptide from the amino end of the heavy chain; this
CC reaction, which occurs at the surface of endothelial cells, is
CC strongly promoted by thrombomodulin.
CC -!- TISSUE SPECIFICITY: Plasma; synthesized in the liver.
CC -!- PTM: The vitamin K-dependent, enzymatic carboxylation of some Glu
CC residues allows the modified protein to bind calcium.
CC -!- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC another site, beyond the GLA domain. This GLA-independent binding
CC site is necessary for the recognition of the
CC thrombin-thrombomodulin complex.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF191307; AAG28380.1;
CC HSP: P4070; 1PCU.
CC MEROPS: S01.218;
CC InterPro: IPR000152; Asx_hydroxyl.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR001881; EGF-Ca.
CC InterPro: IPR001254; Ser_protease_Try.
CC InterPro: IPR000294; VitK_dep_GLA.
CC Pfam: PF00008; EGF; 2.
CC Pfam: PF00089; trypsin; 1.
CC Pfam: PF00594; gla; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00001; GLABLOOD.
CC SMART: SM00181; EGF; 2.
CC SMART: SM00001; EGF_like; 2.
CC SMART: SM00069; GLA; 1.
CC SMART: SM00020; TRYP_SPC; 1.
CC PROSITE: PS00010; ASX_HYDROXYL; 1.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01186; EGF_2; 2.
CC PROSITE: PS01187; EGF_CA; 1.
CC PROSITE: PS00011; GLU_CARBOXYLATION; 1.
CC PROSITE: PS02040; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Glycoprotein; Serine protease;
KW gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.
FT SIGNAL 1
FT PROPEP 19 41
FT CHAIN 42 459
FT CHAIN 42 196
FT CHAIN 199 459
FT PEPTIDE 199 213
FT SITE 213 214
FT ACTIVATION PEPTIDE (BY SIMILARITY).
FT CLEAVAGE (BY THROMBIN) (BY

```
FT DOMAIN 96 131 SIMILARITY).
FT DOMAIN 135 175 EGF-LIKE 1.
FT DOMAIN 214 459 SERINE PROTEASE.
FT MOD_RES 47 47 GAMMA-CARBOXYGLUTAMIC ACID (BY
FT MOD_RES 48 48 SIMILARITY).
FT MOD_RES 55 55 GAMMA-CARBOXYGLUTAMIC ACID (BY
FT MOD_RES 57 57 SIMILARITY).
FT MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID (BY
FT MOD_RES 61 61 SIMILARITY).
FT MOD_RES 66 66 GAMMA-CARBOXYGLUTAMIC ACID (BY
FT MOD_RES 67 67 SIMILARITY).
FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID (BY
FT MOD_RES 112 112 SIMILARITY).
FT ACT_SITE 255 255 HYDROXYLATION (BY SIMILARITY).
FT ACT_SITE 301 301 CHARGE RELAY SYSTEM.
FT ACT_SITE 400 400 CHARGE RELAY SYSTEM.
FT DISULFID 58 63 BY SIMILARITY.
FT DISULFID 91 110 BY SIMILARITY.
FT DISULFID 100 105 BY SIMILARITY.
FT DISULFID 104 119 BY SIMILARITY.
FT DISULFID 121 130 BY SIMILARITY.
FT DISULFID 139 150 BY SIMILARITY.
FT DISULFID 146 159 BY SIMILARITY.
FT DISULFID 161 174 BY SIMILARITY.
FT DISULFID 182 321 INTERCHAIN (BY SIMILARITY).
FT DISULFID 240 256 BY SIMILARITY.
FT DISULFID 371 385 BY SIMILARITY.
FT DISULFID 396 424 BY SIMILARITY.
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 459 AA; 51866 MW; 8541AAC14CC16D09 CRC64;
```

Query Match 89.8%; Score 53; DB 1; Length 459;
Best Local Similarity 90.0%; Pred. No. 0.052;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CEGDSGGPFV 10
Db 396 CEGDSGGPMV 405

Search completed: May 12, 2003, 15:36:37
Job time : 6.8 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:31:22 ; Search time 28.6 Seconds
(without alignments)
46.591 Million cell updates/sec

Title: US-09-909-348-2
Perfect score: 50
Sequence: 1 CXGDSGCPXV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	47	94.0	15	AAW58061	Human prostate spe
2	47	94.0	15	AAW58062	Human prostate spe
3	47	94.0	20	AAW11019	Antigen peptide de
4	47	94.0	214	22 ABB66284	Drosophila melanog
5	47	94.0	237	16 AAR84671	Mature kallikrein
6	47	94.0	237	16 AAR77098	Prostate-specific
7	47	94.0	237	16 AAW83213	Prostate-specific
8	47	94.0	237	19 AAW56086	Human prostate spe
9	47	94.0	237	19 AAW45398	Prostate-specific
10	47	94.0	237	20 AAW96187	Human prostate spe

11	47	94.0	237	21 AAB11041	Human prostate-spe
12	47	94.0	237	22 AAB19818	Prostate specific
13	47	94.0	237	22 AAB19819	Prostate specific
14	47	94.0	240	18 AAW11023	Human prostate spe
15	47	94.0	245	23 AAU98921	Human prostate spe
16	47	94.0	248	22 ABB65295	Drosophila melanog
17	47	94.0	250	17 AAW03130	Prostate-specific
18	47	94.0	252	22 ABB63479	Drosophila melanog
19	47	94.0	252	21 AAB21313	Human PSA. Homo s
20	47	94.0	257	22 AAU06276	Prostate specific
21	47	94.0	259	22 ABB58818	Drosophila melanog
22	47	94.0	259	22 ABB62473	Drosophila melanog
23	47	94.0	260	22 ABB57988	Drosophila melanog
24	47	94.0	261	18 AAW13649	Human prostatic sp
25	47	94.0	261	21 AAB21317	Human PSA. Homo s
26	47	94.0	261	21 AAY77842	Human prostate-spe
27	47	94.0	261	21 AAY56048	Human prostate spe
28	47	94.0	261	22 AAG62144	Human prostate spe
29	47	94.0	261	22 AAG62146	Prostate tumour an
30	47	94.0	261	22 AAB74821	Drosophila melanog
31	47	94.0	262	22 ABB65083	Drosophila melanog
32	47	94.0	266	22 ABB64032	Drosophila melanog
33	47	94.0	271	22 ABB59232	Drosophila melanog
34	47	94.0	276	22 ABB64035	Drosophila melanog
35	47	94.0	287	22 ABB62796	Drosophila melanog
36	47	94.0	296	22 AAB85076	Amino acid sequenc
37	47	94.0	346	21 AAY82701	Tick derived serin
38	47	94.0	362	22 ABB62017	Drosophila melanog
39	47	94.0	375	21 AAB08449	A human prostate s
40	47	94.0	412	22 ABB62250	Drosophila melanog
41	47	94.0	477	22 ABB65796	Drosophila melanog
42	47	94.0	580	22 ABB60547	Drosophila melanog
43	47	94.0	611	22 ABB65794	Drosophila melanog
44	47	94.0	692	22 AAG62154	Human WTI/PSA fusi
45	47	94.0	774	22 ABB70674	Drosophila melanog

ALIGNMENTS

RESULT 1

AAW58061
ID AAW58061 standard; peptide; 15 AA.
XX AC AAW58061;
XX DT 11-AUG-1998 (first entry)
XX DE Human prostate specific antigen peptide SEQ ID NO:79.
XX KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;
XX KW detection; cancer; serine protease.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO9810292-A1.
XX PD 12-MAR-1998.
XX PF 25-AUG-1997; 97WO-US14909.
XX PR 06-SEP-1996; 96US-0025404.
XX (CENZ) CENTOCOR INC.
XX PI Heavner GA;
XX DR WPI; 1998-193789/17.
XX PT Monoclonal antibodies specific for prostate specific antigen -
XX PT useful, e.g. in screening for prostate or breast cancer and
XX PT especially to distinguish between benign prostatic hyperplasia and

PT prostate cancer

XX Example; Page 63; 84pp; English.

XX The present sequence represents a prostate specific antigen (PSA) peptide. Monoclonal antibodies specific for PSA and hybridomas producing them have been developed. The antibodies: (a) bind to free PSA; (b) are monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL) or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and 15.2, 156 or 225 binding to amino acids 139-144 (EELFLTP) and 55-60 (SLFHEP) respectively of free and bound PSA, or fragments. The antibodies are useful to detect PSA. For detecting free PSA only, an immunoassay comprising a solid support with attached labelled monoclonal antibody specific for free PSA (especially (b)) and a PSA standard can be used. To detect both free and bound PSA, a second solid support with attached (differently labelled) monoclonal antibody binding free and bound PSA (especially selected from (c)) can be used either with, or in place of, the first solid support. The antibodies are useful in cancer screening, especially prostate and breast cancer. By obtaining total and free PSA values, their ratio can be used to separate prostatic cancer (PCA) from benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also useful after radical prostatectomy, to predict disease persistence. The antibodies allow evaluation of PSA free/total ratio, enabling separation of BPH and PCA patients with PSA values 4-10 ng/ml not previously possible by total PSA testing, avoiding biopsies.

XX Sequence 15 AA;

Query Match 94.08; Score 47; DB 19; Length 15;
 Best Local Similarity 80.08; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | | | | | | | |
 DB 5 CSGDSGGPLV 14

RESULT 2
 AAW58062
 ID AAW58062 standard; peptide; 15 AA.
 AC AAW58062;
 XX 11-AUG-1998 (first entry)
 DT Human prostate specific antigen peptide SEQ ID NO:80.
 DE Human; prostate specific antigen; PSA; epitope; monoclonal antibody;
 KW detection; cancer; serine protease.
 OS Synthetic.
 OS Homo sapiens.
 XX WO9810292-A1.
 XX 12-MAR-1998.
 XX 25-AUG-1997; 97WO-US14909.
 XX 06-SEP-1996; 96US-0025404.
 XX (CENZ) CENTOCOR INC.
 XX Heavenr GA;
 XX WPI; 1998-193789/17.
 XX Monoclonal antibodies specific for prostate specific antigen -
 PT useful, e.g. in screening for prostate or breast cancer and
 PT especially to distinguish between benign prostatic hyperplasia and
 XX prostate cancer
 PS Example; Page 63; 84pp; English.

XX The present sequence represents a prostate specific antigen (PSA) peptide. Monoclonal antibodies specific for PSA and hybridomas producing them have been developed. The antibodies: (a) bind to free PSA; (b) are monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL) or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and 15.2, 156 or 225 binding to amino acids 139-144 (EELFLTP) and 55-60 (SLFHEP) respectively of free and bound PSA, or fragments. The antibodies are useful to detect PSA. For detecting free PSA only, an immunoassay comprising a solid support with attached labelled monoclonal antibody specific for free PSA (especially (b)) and a PSA standard can be used. To detect both free and bound PSA, a second solid support with attached (differently labelled) monoclonal antibody binding free and bound PSA (especially selected from (c)) can be used either with, or in place of, the first solid support. The antibodies are useful in cancer screening, especially prostate and breast cancer. By obtaining total and free PSA values, their ratio can be used to separate prostatic cancer (PCA) from benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also useful after radical prostatectomy, to predict disease persistence. The antibodies allow evaluation of PSA free/total ratio, enabling separation of BPH and PCA patients with PSA values 4-10 ng/ml not previously possible by total PSA testing, avoiding biopsies.

XX Sequence 15 AA;

Query Match 94.08; Score 47; DB 19; Length 15;
 Best Local Similarity 80.08; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | | | | | | | |
 DB 2 CSGDSGGPLV 11

RESULT 3
 AAW11019
 ID AAW11019 standard; peptide; 20 AA.
 AC AAW11019;
 XX 27-OCT-1997 (first entry)
 DT Antigen peptide derived from human prostate specific antigen.
 DE PSA; kallikrein; prostate cancer; antigenic; polyclonal antisera.
 KW Homo sapiens.
 OS WO9640754-A1.
 XX 19-DEC-1996.
 XX 06-JUN-1996; 96WO-US09303.
 XX 07-JUN-1995; 95US-0472228.
 XX (TEXA) UNIV TEXAS SYSTEM.
 XX Fritsche HA, Johnston DA, Kokolus WJ;
 XX WPI; 1997-108633/10.
 XX Antigen peptide derived from prostate specific antigen - does not
 PT cross react with related kallikreins, for diagnosis of prostate
 PT cancer
 XX Example 2; Page 42; 74pp; English.
 XX The present sequence represents a novel peptide which has 20 contiguous
 CC amino acids derived from the 240 residue sequence of the human prostate
 CC specific antigen (PSA) (see AAW11023). The preferred peptide has two
 CC hydrophobic regions and one hydrophilic region each of about 5 aa in
 CC length, arranged as follows: hydrophobic-hydrophilic-hydrophobic.

CC Polyclonal antisera specific for the antigenic peptide may be used in a
 CC method for diagnosing prostate cancer in vivo or in vitro. The peptide
 CC represents a sequence unique to PSA which does not cross react with
 CC certain kallikreins.

XX SQ Sequence 20 AA;

Query Match 94.0%; Score 47; DB 18; Length 20;
 Best Local Similarity 80.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | | | | | | | |
 Db 4 CSGDSGGPLV 13

RESULT 4

ABB66284
 ID ABB66284 standard; Protein; 214 AA.

XX AC ABB66284;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 25644.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX OS Drosophila melanogaster.

XX FN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX FI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL10387.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 25644; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 214 AA;

Query Match 94.0%; Score 47; DB 22; Length 214;
 Best Local Similarity 80.0%; Pred. No. 77;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | | | | | | | |

Db 160 CSGDSGGPLV 169

RESULT 5

AAR84671
 ID AAR84671 standard; Protein; 237 AA.

XX AC AAR84671;

XX DT 04-MAR-1996 (first entry)

XX DE Mature kallikrein HK3.

XX KW Kallikrein HK3; serine protease; prostate-specific antigen;
 KW prostate cancer; diagnosis.

XX OS Homo sapiens.

XX FN WO9530758-A1.

XX PD 16-NOV-1995.

XX PF 09-MAY-1995; 95WO-US06157.

XX PR 02-MAY-1995; 95US-0427767.

XX PR 10-MAY-1994; 94US-0241174.

XX PA (HYBR-) HYBRITECH INC.

XX PA (MAYO-) MAYO FOUNDATION.

XX PI Saedi MS, Tindall DJ, Young CYF;

XX DR WPI; 1995-404123/51.

PT New isolated prostate-specific kallikrein polypeptide(s) - used to
 PT develop prods. for use in assays for such polypeptide(s), partic.
 PT for diagnosis and monitoring of prostate cancer

XX PS Disclosure; Page 31; 61pp; English.

XX Human kallikrein HK3 (AAR84671) is a prostate-specific kallikrein
 CC showing homology to prostate-specific glandular kallikrein HK2
 CC (AAR84667). Nonhomologous regions of the kallikreins can be used
 CC for prepn. of antibodies specific to HK2.

XX SQ Sequence 237 AA;

Query Match 94.0%; Score 47; DB 16; Length 237;
 Best Local Similarity 80.0%; Pred. No. 84;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | | | | | | | |

Db 185 CSGDSGGPLV 194

RESULT 6

AAR77098

ID AAR77098 standard; Protein; 237 AA.

XX AC AAR77098;

XX DT 16-JAN-1996 (first entry)

XX DE Prostate-specific antigen.

XX KW Prostate-specific antigen; prostate cancer; diagnosis.

XX OS Homo sapiens.

XX PN WO9528498-A1.

XX PD 26-OCT-1995.

XX PF 14-APR-1995; 95WO-US04680.
 XX XX 15-APR-1994; 94US-0229391.
 XX XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX PI Buttyan R, Katz AE, Olsson CA, Raffo A;
 XX DR WPI; 1995-373812/48.
 XX DR N-PSDB; AAT04864.
 XX PT Method for enhancing prostate-specific antigen detection - provides
 XX PT sensitive means to identify early stages of prostate cancer.
 XX PS Disclosure; Page 69-70; 94pp; English.
 XX CC A prostate-specific antigen (PSA) is detected by isolating mRNA from
 XX CC a sample, producing cDNA, and subjecting the cDNA to RT-PCR using
 XX CC primers specific for the human PSA coding region (given in AAT04864).
 XX SQ Sequence 237 AA;

Query Match 94.0%; Score 47; DB 16; Length 237;
 Best Local Similarity 80.0%; Pred. No. 84;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGGPXV 10
 | | | | | | | |
 Db 185 CSGDSGGGPLV 194

RESULT 7
 AAW83213
 ID AAW83213 standard; Protein; 237 AA.

XX AC AAW83213;
 XX DT 09-FEB-1999 (first entry)
 XX DE Prostate-specific antigen protein hk3.

XX KW Prostate cancer; detection; hk2; hk1; hk3; phk2; tissue kallikrein;
 XX KW phk2; prostate-specific glandular kallikrein protein; PSA; human;
 XX KW prostate-specific antigen.
 XX OS Homo sapiens.

XX PN WO9846795-A1.
 XX PD 22-OCT-1998.
 XX PF 09-APR-1998; 98WO-US07037.

XX PR 11-APR-1997; 97US-0843076.
 XX PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX PA (MAYO-) MAYO FOUNDATION.

XX PI Slawin KM, Tindall DJ, Young CYF;
 XX DR WPI; 1998-594592/50.
 XX PT Detection of human kallikrein 2 RNA - by reverse transcription and
 XX PT amplification by PCR, for detecting, monitoring and staging of
 XX PT prostate cancer

XX PS Disclosure; Fig 1; 90pp; English.
 XX CC The present invention describes a diagnostic method for detecting human
 XX CC kallikrein 2 (hk2) DNA. The method comprises: (a) contacting DNA obtained
 XX CC by reverse transcription (RT) of RNA from a human physiological sample
 XX CC which comprises cells suspected of containing hk2 RNA with at least 2

CC oligonucleotides to amplify the DNA by PCR to yield amplified hk2 DNA,
 CC where the conditions amplify the DNA obtained by RT of RNA from at least
 CC one cell containing hk2 in a sample which comprises at least 107 to 109
 CC cells; and (b) detecting the presence of the amplified hk2 DNA. The
 CC method can be used for detecting, monitoring the progression of and
 CC pathologically staging prostate cancer. The present sequence represents
 CC prostate-specific antigen protein hk3.

XX SQ Sequence 237 AA;

Query Match 94.0%; Score 47; DB 19; Length 237;
 Best Local Similarity 80.0%; Pred. No. 84;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGGPXV 10
 | | | | | | | |
 Db 185 CSGDSGGGPLV 194

RESULT 8
 AAW56086
 ID AAW56086 standard; Protein; 237 AA.

XX AC AAW56086;
 XX DT 11-AUG-1998 (first entry)
 XX DE Human prostate specific antigen protein.

XX KW Human; prostate specific antigen; PSA; epitope; monoclonal antibody;
 XX KW detection; cancer; serine protease.
 XX OS Homo sapiens.
 XX PN WO9810292-A1.
 XX PD 12-MAR-1998.
 XX PF 25-AUG-1997; 97WO-US14909.
 XX PR 06-SEP-1996; 96US-0025404.

XX PA (CENZ) CENTOCOR INC.
 XX PI Heavner GA;
 XX DR WPI; 1998-193789/17.

XX PT Monoclonal antibodies specific for prostate specific antigen -
 XX PT useful, e.g. in screening for prostate or breast cancer and
 XX PT especially to distinguish between benign prostatic hyperplasia and
 XX PT prostate cancer
 XX PS Disclosure; Page 32-33; 84pp; English.
 XX CC The present sequence represents prostate specific antigen (PSA).
 XX CC Monoclonal antibodies specific for PSA and hybridomas producing them
 XX CC have been developed. The antibodies: (a) bind to free PSA; (b) are
 XX CC monoclonal antibody 365 binding to amino acids 82-87 of free PSA (LKNRFL)
 XX CC or fragment, and (c) are monoclonal antibodies 10, 11, 16 or 22.2 and
 XX CC 15.2, 156 or 225 binding to amino acids 139-144 (EELFLTP) and 55-60
 XX CC (SIFHPE) respectively of free and bound PSA, or fragments. The antibodies
 XX CC are useful to detect PSA. For detecting free PSA only, an immunoassay
 XX CC comprising a solid support with attached labelled monoclonal antibody
 XX CC specific for free PSA (especially (b)) and a PSA standard can be used.
 XX CC To detect both free and bound PSA, a second solid support with attached
 XX CC (differently labelled) monoclonal antibody binding free and bound PSA
 XX CC (especially selected from (c)) can be used either with, or in place of,
 XX CC the first solid support. The antibodies are useful in cancer screening,
 XX CC especially prostate and breast cancer. By obtaining total and free PSA
 XX CC values, their ratio can be used to separate prostatic cancer (PCa) from
 XX CC benign prostatic hyperplasia (BHP) patients. Measurement of PSA is also
 XX CC useful after radical prostatectomy, to predict disease persistence. The

CC antibodies allow evaluation of PSA free/total ratio, enabling separation
 CC of BPH and PCA patients with PSA values 4-10 ng/ml not previously
 CC possible by total PSA testing, avoiding biopsies.

XX SQ Sequence 237 AA;

Query Match 94.0%; Score 47; DB 19; Length 237;

Best Local Similarity 80.0%; Pred. No. 84;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | | | | | | | |
 Db 185 CSGDSGGPLV 194

RESULT 9

AAW45398
 ID AAW45398 standard; Protein; 237 AA.

XX AC AAW45398;

XX DT 06-JUL-1998 (first entry)

XX DE Prostate-specific antigen protein hK3 (PSA).

XX KW Prostate-specific antigen; hK3; PSA; human; hK2;

XX KW prostate-specific glandular kallikrein; prostate carcinoma;

XX KW prostate cancer; benign prostatic hyperplasia; diagnosis.

XX OS Homo sapiens.

XX PN W09802748-A1.

XX PD 22-JAN-1998.

XX PE 15-JUL-1997; 97WO-US12322.

XX PR 15-JUL-1996; 96US-0680869.

XX PA (HYBR-) HYBRITECH INC.

XX PA (MAYO-) MAYO FOUNDATION.

XX PI Grauer L, Klee GG, Mikolajczyk SD, Saedi M, Tindall DJ;

PI Young CYF;

XX DR WPI; 1998-120378/11.

XX PT Diagnostic methods using antibodies which bind prostate antigens -
 PT useful for, e.g. monitoring treatment or progression of prostate
 PT cancer

PS Disclosure; Page 55-56; 100pp; English.

CC This polypeptide comprises human prostate-specific antigen (PSA)

CC hK3. PSA levels are widely used as a prognostic indicator of

CC prostate carcinoma. The invention provides a novel diagnostic

CC method comprising contacting antibodies that specifically bind to

CC human prostate-specific glandular kallikrein protein hK2 (see

CC W453950 or pro-hK2 (phK2, see AAW45396), but not with PSA, with a

CC sample of physiological fluid from a human. The assay is based on

CC the discovery that phK2 is detected in the supernatant of a

CC prostate carcinoma cell line and that hK2 is present in human

CC physiological fluid from prostate cancer cells. The method is

CC useful for monitoring the treatment and/or progression of prostate

CC cancer, or for the early detection of prostate cancer in males that

CC have BPH or a high grade prostatic neoplasia (HPGN) or whose family

CC members have or had BPH, HPGN or a prostate cancer. Monitoring the

CC presence and/or amount of hK2 complexes with plasma proteins may

CC also be important in distinguishing between prostate cancer and BPH.

XX SQ Sequence 237 AA;

Query Match 94.0%; Score 47; DB 19; Length 237;

Best Local Similarity 80.0%; Pred. No. 84;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | | | | | | | |
 Db 185 CSGDSGGPLV 194

RESULT 10

AAW96187
 ID AAW96187 standard; Protein; 237 AA.

XX AC AAW96187;

XX DT 27-APR-1999 (first entry)

XX DE Human prostate specific antigen (PSA).

XX KW Human kallikrein 2; hK2; breast cancer; diagnosis; detection;

XX KW treatment; monitoring; prostate specific antigen; PSA.

XX OS Homo sapiens.

XX PN W09859073-A1.

XX PD 30-DEC-1998.

XX PF 19-JUN-1998; 98WO-US12840.

XX PR 20-JUN-1997; 97US-0050963.

XX PA (KLEE/) KLEE G G.

XX PA (MAYO-) MAYO FOUNDATION.

XX PA (TIND/) TINDALL D J.

XX PA (YOUN/) YOUNG C Y F.

XX PI Klee GG, Tindall DJ, Young CYF;

XX DR WPI; 1999-105632/09.

XX DR N-PSDB; AAX08947.

XX PT Use of human kallikrein 2 - as a marker for developing products for
 PT the diagnosis, prognosis, monitoring and treatment of breast cancer

XX PS Disclosure; Page 33-34; 67pp; English.

XX Human kallikrein 2 (hK2) is expressed at elevated levels relative to

CC the prostate cancer antigen, prostate specific antigen (PSA) by

CC breast cancer cells when stimulated by androgens. Detecting levels

CC of hK2 may provide a simple diagnostic tool for detecting or

CC determining breast cancer. Detecting hK2 is achieved by producing

CC an hK2 DNA from hK2 RNA by reverse transcription. The hK2 DNA

CC obtained is then amplified by PCR and detected using probes.

CC Determination of breast cancer may also be determined by exposing

CC the hK2 polypeptide to an agent which binds to it and then detecting

CC the binary complex formed. The amount of complex formed correlates

CC directly to the presence of breast cancer in the human individual.

CC The methods can be used in the diagnosis, treatment and/or

CC monitoring of the progression or remission of breast cancer and/or

CC monitoring hK2 levels.

XX SQ Sequence 237 AA;

Query Match 94.0%; Score 47; DB 20; Length 237;

Best Local Similarity 80.0%; Pred. No. 84;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | | | | | | | |
 Db 185 CSGDSGGPLV 194

RESULT 11

AAB11041
ID AAB11041 standard; peptide; 237 AA.
XX AC AAB11041;
XX DT 08-FEB-2001 (first entry)
XX XX Human prostate-specific antigen N-terminal fragment #2.
XX DE
XX XX Prostate-specific antigen; PSA; human; detection; prostatic cancer;
KW KW Isoform-specific monoclonal antibodies; marker.
XX OS Homo sapiens.
XX PN EF1043394-A2.
XX PD 11-OCT-2000.
XX XX 08-APR-2000; 2000EP-0107649.
XX PF 10-APR-1999; 99DE-1016242.
XX PR (SERA-) SERATEC GES BIOTECHNOLOGIE MBH.
XX PA Volk M, Mast W;
XX PI WPI; 2000-674036/66.
XX DR
XX XX Composition containing isoforms of prostate-specific antigen, useful as
PT a reference standard for analysis and for producing isoform-specific
PT antibodies -
XX XX
XX PS Disclosure; Figure 4; 23pp; German.
XX CC This invention describes a novel composition (A) having a total
CC concentration of prostate-specific antigen (PSA). (A) is useful as a
CC reference in methods for detecting PSA (a marker for prostatic cancer) in
CC human blood or serum and to generate isoform-specific monoclonal
CC antibodies (MAb) against PSA, for use in immunological assays. (A)
CC contains all the isoforms of PSA and can be isolated without significant
CC formation of artefacts by autolysis. Antibodies specific for
CC particular PSA isoforms can improve diagnostic specificity.
XX XX
XX SQ Sequence 237 AA;
Query Match 94.0%; Score 47; DB 21; Length 237;
Best Local Similarity 80.0%; Pred. No. 84;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXGDSGGPXV 10
Db 185 CSGDSGGPLV 194
RESULT 12
AAB19818
ID AAB19818 standard; protein; 237 AA.
XX AC AAB19818;
XX DT 05-MAR-2001 (first entry)
XX DE Prostate specific antigen elevated in benign prostatic hyperplasia.
XX KW Prostate specific antigen; PSA; benign prostatic hyperplasia; BPH;
KW marker; immunoassay; diagnosis; human.
XX OS Homo sapiens.
XX XX
XX Key Location/Qualifiers
FT Cleavage-site 1
FT /note- "site of internal peptide bond cleavage"
FT 145
FT Cleavage-site 145

FT Cleavage-site 146
FT /note- "site of internal peptide bond cleavage"
FT 182
FT Cleavage-site 182
FT /note- "site of internal peptide bond cleavage"
PN WO200066718-A1.
XX XX
XX PD 09-NOV-2000.
XX PF 07-APR-2000; 2000WO-US09415.
XX PR 30-APR-1999; 99US-0303208.
XX XX (HYBR-) HYBRITECH INC.
XX PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX XX Mikolajczyk S, Wang T, Rittenhouse H, Wolfert R, Slawin K;
XX PI WPI; 2001-015979/02.
XX DR
XX XX Novel prostate specific antigen (PSA) for distinguishing benign
PT prostatic hyperplasia from prostate cancer, comprises a clip at lysine
PT 182 of the amino acid sequence of a mature form of PSA -
XX XX
XX PS Disclosure; Fig 14; 67pp; English.
XX CC The present sequence represents a mature form of a novel form of
CC prostate specific antigen (PSA). This form of PSA has at least 1
CC clip at Lys-182 and may additionally have 1 or more clips at Lys-1,
CC Lys-145 and Lys-146. Preferably, the PSA contains at least 2 clips
CC at Lys-145 and Lys-182. The novel forms of PSA exist at an
CC elevated level in patients suspected of having benign prostatic
CC hyperplasia (BPH) and therefore may be used as a serum marker or
CC an immunohistological marker to help distinguish BPH from prostate
CC cancer. Antibodies recognising the novel forms of PSA and
CC immunoassays that detect and determine the novel forms of PSA are
CC also provided.
XX XX
XX SQ Sequence 237 AA;
Query Match 94.0%; Score 47; DB 22; Length 237;
Best Local Similarity 80.0%; Pred. No. 84;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXGDSGGPXV 10
Db 185 CSGDSGGPLV 194
RESULT 13
AAB19819
ID AAB19819 standard; protein; 237 AA.
XX AC AAB19819;
XX DT 05-MAR-2001 (first entry)
XX DE Prostate specific antigen specific to benign prostatic hyperplasia.
XX KW Prostate specific antigen; PSA; BPSA; benign prostatic hyperplasia;
KW BPH; marker; diagnosis; human.
XX OS Homo sapiens.
XX XX
XX Key Location/Qualifiers
FT Cleavage-site 1
FT /note- "site of internal peptide bond cleavage"
FT 145
FT Cleavage-site 145
FT /note- "site of internal peptide bond cleavage"
FT 146
FT Cleavage-site 146
FT /note- "site of internal peptide bond cleavage"
FT 182
FT Cleavage-site 182

FT /note= "site of internal peptide bond cleavage"

PN WO200067030-A1.

XX

PD 09-NOV-2000.

XX

PF 12-APR-2000; 2000WO-US09789.

XX

XX 30-APR-1999; 99US-0303339.

XX (HYBR-) HYBRITECH INC.

PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX

PI Mikolajczyk S, Wang T, Rittenhouse H, Wolfert R, Slawin K;

XX

DR WPI; 2001-007264/01.

XX

PT Distinguishing prostate cancer from benign prostatic hyperplasia using

PT different forms of prostate specific antigen contained in a sample

PT involves mathematically combining amounts of ProPSA and BPSA

XX

PS Disclosure; Fig 14; 60pp; English.

XX

CC The present sequence represents a mature form of a novel form of

CC prostate specific antigen (PSA). This form of PSA has at least 1

CC clip at Lys-182 and may additionally have 1 or more clips at Lys-1,

CC Lys-145 and Lys-146. BPSA, a form of PSA that comprises at least 1

CC clip at Lys-182, exists at an elevated level in the transition zone

CC of benign prostatic hyperplasia (BPH) tissue, compared to

CC peripheral zone cancer and non-cancer prostate tissues. A method

CC for distinguishing prostate cancer from BPH involves determining

CC the ratio of proPSA and BPSA in a sample.

XX

SQ Sequence 237 AA;

Query Match 94.0%; Score 47; DB 22; Length 237;

Best Local Similarity 80.0%; Pred. No. 84;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10

I | I I I I I I I

DB 185 CSGDSGGPLV 194

RESULT 14

AAW11023

ID AAW11023 standard; Protein; 240 AA.

XX

AC AAW11023;

XX

DT 27-OCT-1997 (first entry)

XX

DE Human prostate specific antigen.

XX

KW PSA; kallikrein; prostate cancer; antigenic; polyclonal antisera.

XX

OS Homo sapiens.

XX

PN WO9640754-A1.

XX

PD 19-DEC-1996.

XX

PF 06-JUN-1996; 96WO-US09303.

XX

PR 07-JUN-1995; 95US-0472228.

XX

XX (TEXA) UNIV TEXAS SYSTEM.

XX

PI Fritsche HA, Johnston DA, Kokolus WJ;

XX

DR WPI; 1997-108633/10.

XX

PT Antigen peptide derived from prostate specific antigen - does not

PT cross react with related kallikreins, for diagnosis of prostate

PT cancer

XX

PS Claim 1; Fig 1; 74pp; English.

XX

CC The present sequence represents human prostate specific antigen (PSA)

CC which is used to produce novel peptides derived from the 240 residue

CC sequence. The preferred peptides have two hydrophobic regions and one

CC hydrophilic region each of about 5 aa in length, arranged as follows:

CC hydrophobic-hydrophilic-hydrophobic. Polyclonal antisera specific for

CC the antigenic peptides may be used in a method for diagnosing prostate

CC cancer in vivo or in vitro. The peptides represent a sequence unique

CC to PSA which does not cross react with certain kallikreins.

XX

SQ Sequence 240 AA;

Query Match 94.0%; Score 47; DB 18; Length 240;

Best Local Similarity 80.0%; Pred. No. 84;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10

I | I I I I I I I

DB 188 CSGDSGGPLV 197

RESULT 15

AAU98921

ID AAU98921 standard; Protein; 245 AA.

XX

AC AAU98921;

XX

DT 24-SEP-2002 (first entry)

XX

DE Human prostate specific antigen (PSA) variant.

XX

KW Human; prostate specific antigen; PSA; cytostatic; antiviral;

XX immunostimulant; cell-mediated immune response; tumour; breast cancer;

XX virus infection; prostate cancer; colorectal cancer; pancreatic cancer;

XX lymphoma; leukaemia; hepatitis virus; lentivirus; herpesvirus;

XX human immunodeficiency virus; HIV; flavivirus; pestivirus.

XX

OS Homo sapiens.

XX

PH Key

FT Region

FT 44..750

FT /note= "This region is specifically claimed"

FT Region

FT 55..750

FT /note= "This region is specifically claimed"

FT Region

FT 61..750

FT /note= "This region is specifically claimed"

FT Region

FT 76..750

FT /note= "This region is specifically claimed"

FT Region

FT 94..750

FT /note= "This region is specifically claimed"

XX

PN WO200240059-A2.

XX

PD 23-MAY-2002.

XX

PF 01-NOV-2001; 2001WO-US45626.

XX

PR 01-NOV-2000; 2000US-0704232.

XX

XX (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.

PA (MINC/) MINCHEFF M S.

PA (LOUK/) LOUKINOV D I.

PA (ZOUT/) ZOUTBAK S.

XX

PI Mincheff MS, Loukinov DI, Zoubak S;

XX

DR WPI; 2002-527524/56.

XX

DR N-PSDB; ABR86205.

XX

PT Inducing a cell-mediated immune response against a target antigen,
PT reducing undesired cells and stimulating presentation of an antigen by
PT a cell, comprises administering a polynucleotide encoding a variant of
an antigen -

XX
XX
PS Claim 15; Page 122-123; 146pp; English.

XX
CC The invention relates to a method of inducing a cell-mediated immune
CC response against a cell comprising a target antigen (I) in a subject,
CC treating a subject having undesired cells, for example tumour cells
CC or virally infected cells (C), reducing the number of (C) in a subject,
CC and stimulating presentation of (I) by a cell. This is done by
CC administering a polynucleotide (II) encoding a variant of (I), so that
CC (II) expressed in a cell and cell-mediated immune response is induced.
CC The method can be used to treat prostate cancer, breast cancer,
CC colorectal cancer and pancreatic cancer, as well as lymphomas and
CC leukaemias. The method is also useful in treating chronic viral
CC infections such as those caused by hepatitis viruses, lentiviruses
CC (including human immunodeficiency virus (HIV)), herpesviruses and the
CC flaviviruses and pestiviruses. The present sequence represents the
CC human prostate specific antigen (PSA) variant which has the signal
CC sequence deleted, used as a target antigen in the method of the
CC invention.

XX
SQ Sequence 245 AA;

Query Match 94.0%; Score 47; DB 23; Length 245;

Best Local Similarity 80.0%; Pred. No. 86;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10

| | | | | | |

Db 193 CSGDSGGPLV 202

Search completed: May 12, 2003, 15:36:03
Job time : 29.6 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein -- protein search, using sw model

Run on: May 12, 2003, 15:33:37 ; Search time 9.6 Seconds
(without alignments)
30.649 Million cell updates/sec

Title: US-09-909-348-2

Perfect score: 50

Sequence: 1 CXGDSGGPXV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	94.0	20	1	US-08-472-228A-19
2	47	94.0	20	4	US-09-146-831-19
3	47	94.0	20	5	PCT-US96-09303-19
4	47	94.0	237	1	US-08-096-946-11
5	47	94.0	237	2	US-08-844-024-2
6	47	94.0	237	2	US-08-718-547-2
7	47	94.0	237	3	US-08-768-859A-1
8	47	94.0	237	3	US-08-767-820A-1
9	47	94.0	237	3	US-08-622-046B-7
10	47	94.0	237	4	US-08-944-483-38
11	47	94.0	237	4	US-09-100-284-3
12	47	94.0	237	4	US-09-303-339-2
13	47	94.0	237	5	PCT-US94-07329-11
14	47	94.0	237	5	PCT-US95-06157-1
15	47	94.0	240	1	US-08-472-228A-1
16	47	94.0	240	4	US-09-146-831-1
17	47	94.0	240	5	PCT-US96-09303-1
18	47	94.0	251	4	US-08-944-483-47
19	47	94.0	261	1	US-08-744-026-5
20	47	94.0	261	2	US-09-102-732-5
21	47	94.0	261	3	US-09-083-521-6
22	47	94.0	261	4	US-09-261-767-5
23	47	94.0	261	4	US-09-413-049-1
24	47	94.0	262	3	US-09-025-059-4
25	47	94.0	1019	1	US-08-296-014A-4
26	47	94.0	1019	2	US-08-596-405-4
27	47	94.0	1019	2	US-08-877-620-4

28 47 94.0 1083 1 US-08-296-014A-2 Sequence 2, Appli
29 47 94.0 1083 2 US-08-596-405-2 Sequence 2, Appli
30 47 94.0 1083 2 US-08-877-620-2 Sequence 2, Appli
31 46 92.0 12 6 5352864-3 Patent No. 5352864
32 46 92.0 23 4 US-08-981-088-2 Sequence 2, Appli
33 46 92.0 23 4 US-08-330-594-8 Sequence 8, Appli
34 46 92.0 68 1 US-08-485-455D-77 Sequence 77, Appli
35 46 92.0 68 2 US-08-482-130C-77 Sequence 77, Appli
36 46 92.0 68 2 US-08-484-211C-77 Sequence 77, Appli
37 46 92.0 68 3 US-08-906-769-77 Sequence 77, Appli
38 46 92.0 68 3 US-08-906-616-77 Sequence 77, Appli
39 46 92.0 68 4 US-08-817-795-77 Sequence 77, Appli
40 46 92.0 68 4 US-08-485-443B-77 Sequence 77, Appli
41 46 92.0 68 4 US-08-639-075A-77 Sequence 77, Appli
42 46 92.0 68 4 US-09-012-431-77 Sequence 77, Appli
43 46 92.0 68 4 US-09-012-692-77 Sequence 77, Appli
44 46 92.0 68 4 US-08-906-613-77 Sequence 77, Appli
45 46 92.0 68 5 PCT-US95-14442A-77 Sequence 77, Appli

ALIGNMENTS

RESULT 1

US-08-472-228A-19
; Sequence 19, Application US/08472228A
; Patent No. 5807978
; GENERAL INFORMATION:
; APPLICANT: Kokolus, William J.
; APPLICANT: Fritsche, Herbert A.
; APPLICANT: Johnston, Dennis A.
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES OF PROSTATE
; TITLE OF INVENTION: SPECIFIC ANTIGEN
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,228A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:388
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-472-228A-19

Query Match 94.0%; Score 47; DB 1; Length 20;
Best Local Similarity 80.0%; Pred. No. 0.12;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CXGDSGGPXV 10
| | | | | | | |

Db 4 CSGDSGGPLV 13

RESULT 2

US-09-146-831-19
; Sequence 19, Application US/09146831
; Patent No. 6326471
; GENERAL INFORMATION:
; APPLICANT: Kokolus, William J.
; APPLICANT: Fritsche, Herbert A.
; APPLICANT: Johnston, Dennis A.
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES OF PROSTATE
; TITLE OF INVENTION: SPECIFIC ANTIGEN
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,831
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/472,228
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:388
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-146-831-19

Query Match 94.0%; Score 47; DB 4; Length 20;
Best Local Similarity 80.0%; Pred. No. 0.12;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10

Db 4 CSGDSGGPLV 13

RESULT 3

PCT-US96-09303-19
; Sequence 19, Application PC/TUS9609303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES OF PROSTATE
; TITLE OF INVENTION: SPECIFIC ANTIGEN
; NUMBER OF SEQUENCES: 22
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/09303
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/472,228
; FILING DATE: 07-JUN-1995
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-09303-19

Query Match 94.0%; Score 47; DB 5; Length 20;
Best Local Similarity 80.0%; Pred. No. 0.12;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10

Db 4 CSGDSGGPLV 13

RESULT 4

US-08-096-946-11
; Sequence 11, Application US/08096946
; Patent No. 5516639
; GENERAL INFORMATION:
; APPLICANT: Tindall, Donald J.
; APPLICANT: Young, Charles Y-F
; APPLICANT: Klee, George G.
; TITLE OF INVENTION: Antibodies Specific for Human
; TITLE OF INVENTION: Prostate Glandular Kallikrein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 5516639west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,946
; FILING DATE: 19930722
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 1016.62-US-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-8081
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Amino acid sequence of PSA (hk3)
US-08-096-946-11

Query Match 94.0%; Score 47; DB 1; Length 237;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10

Db 185 CSGDSGGPLV 194
| | | | | | | |

RESULT 5

US-08-844-024-2
; Sequence 2, Application US/08844024
; Patent No. 5840494
; GENERAL INFORMATION:
; APPLICANT: Katz, Aaron E., et al.
; TITLE OF INVENTION: A Method For Molecular Staging Of
; PROSTATE CANCER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/844,024
; APPLICATION NUMBER: 08/229,391
; FILING DATE: 15-APR-1994
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 2:
; MOLECULE TYPE: protein
; LENGTH: 237 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; QUERY MATCH 94.0%; Score 47; DB 2; Length 237;
; Best Local Similarity 80.0%; Pred. No. 1.1;
; Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10
| | | | | | | |
Db 185 CSGDSGGPLV 194

RESULT 6

US-08-718-547-2
; Sequence 2, Application US/08718547
; Patent No. 5976794
; GENERAL INFORMATION:
; APPLICANT: Katz, Aaron E.; Buttyan, Ralph; Raffo, Anthony;
; APPLICANT: Olsson, Carl A.
; TITLE OF INVENTION: A Method For Molecular Staging Of
; PROSTATE CANCER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.

ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/718,547
; APPLICATION NUMBER: 08/718,547
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 2:
; MOLECULE TYPE: protein
; LENGTH: 237 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; QUERY MATCH 94.0%; Score 47; DB 2; Length 237;
; Best Local Similarity 80.0%; Pred. No. 1.1;
; Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10
| | | | | | | |
Db 185 CSGDSGGPLV 194

RESULT 7

US-08-768-859A-1
; Sequence 1, Application US/08768859A
; Patent No. 6013471
; GENERAL INFORMATION:
; APPLICANT: Tindall, Donald J.
; APPLICANT: Young, Charles Y.F.
; APPLICANT: Saedi, Mohammed S.
; TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/768,859A
; FILING DATE: 17-DECEMBER-1995
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; MOLECULE TYPE: single
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

Query Match 94.0%; Score 47; DB 2; Length 237;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10
| | | | | | | |
Db 185 CSGDSGGPLV 194

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-768-859A-1

Query Match 94.0%; Score 47; DB 3; Length 237;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
Db 185 CSGDSGGPLV 194

RESULT 8

US-08-767-820A-1
Sequence 1, Application US/08767820A
Patent No. 6093796
GENERAL INFORMATION:
APPLICANT: Tindall, Donald J.
APPLICANT: Young, Charles Y.F.
APPLICANT: Saeidi, Mohammed S.
TITLE OF INVENTION: DNA Encoding HK2 Variant Polypeptides
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,820A
FILING DATE: 17-DECEMBER-1996
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 545.002US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-767-820A-1

Query Match 94.0%; Score 47; DB 3; Length 237;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
Db 185 CSGDSGGPLV 194

RESULT 9

US-08-622-046B-7
Sequence 7, Application US/08622046B
Patent No. 6103237
GENERAL INFORMATION:
APPLICANT: Saeidi, Mohammed S.
APPLICANT: Mikolajczyk, Stephen D.
TITLE OF INVENTION: Stable Variant HK2 Polypeptide
NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/622,046B
FILING DATE: 26-MARCH-1996
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 476.001US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6903
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-622-046B-7

Query Match 94.0%; Score 47; DB 3; Length 237;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
Db 185 CSGDSGGPLV 194

RESULT 10

US-08-944-483-38
Sequence 38, Application US/08944483
Patent No. 6232456
GENERAL INFORMATION:
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6232456e
US-08-944-483-38

Query Match 94.0%; Score 47; DB 4; Length 237;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPVX 10
| | | | | | | |
Db 185 CSGDSGGPLV 194

RESULT 11
US-09-100-264-3
Sequence 3, Application US/09100264
Patent No. 6235486
GENERAL INFORMATION:
APPLICANT: Young, Charles Y.F.
APPLICANT: Tindall, Donald J.
APPLICANT: Klee, George G.
TITLE OF INVENTION: Method for detection of breast cancer
FILE REFERENCE: 545,006US1
CURRENT APPLICATION NUMBER: US/09/100,264
CURRENT FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: US 60/050,963
EARLIER FILING DATE: 1997-06-20
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 237
TYPE: PRT
ORGANISM: Homo sapiens
US-09-100-264-3

Query Match 94.0%; Score 47; DB 4; Length 237;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPVX 10
| | | | | | | |
Db 185 CSGDSGGPLV 194

RESULT 12
US-09-303-339-2
Sequence 2, Application US/09303339
Patent No. 6423503
GENERAL INFORMATION:
APPLICANT: Mikolajczyk, Stephen
APPLICANT: Wang, Tang
APPLICANT: Harry, Rittenhouse
APPLICANT: Robert, Wolfert
TITLE OF INVENTION: No. 6423503el Forms of Free Prostate-Specific Antigen (PSA) an
TITLE OF INVENTION: Association with Prostate Tissues From Prostate Peripheral Zone

FILE REFERENCE: 451D-1875
CURRENT APPLICATION NUMBER: US/09/303,339
CURRENT FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 237
TYPE: PRT
ORGANISM: Homo sapiens
US-09-303-339-2

Query Match 94.0%; Score 47; DB 4; Length 237;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPVX 10
| | | | | | | |
Db 185 CSGDSGGPLV 194

RESULT 13
PCT-US94-07329-11
Sequence 11, Application PC/TUS9407329
GENERAL INFORMATION:
APPLICANT: Mayo Foundation for Medical
APPLICANT: Education and Research
TITLE OF INVENTION: Antibodies Specific for Human
TITLE OF INVENTION: Prostate Glandular Kallikrein
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg & Woessner
STREET: 3500 IDS Center
STREET: 80 South Eighth Street
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07329
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Raasch, Kevin W.
REGISTRATION NUMBER: 35,651
REFERENCE/DOCKET NUMBER: 150.62WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Amino acid sequence of PSA (hk3)
PCT-US94-07329-11

Query Match 94.0%; Score 47; DB 5; Length 237;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPVX 10
| | | | | | | |
Db 185 CSGDSGGPLV 194

RESULT 14

PCT-US95-06157-1
; Sequence 1, Application PC/TUS9506157
; GENERAL INFORMATION:
; APPLICANT: Mayo Foundation for Medical Education
; APPLICANT: and Research
; APPLICANT: Hybritech Incorporated
; APPLICANT: Tindall, Donald J.
; APPLICANT: Young, Charles Y.F.
; APPLICANT: Saedi, Mohammed S.
; TITLE OF INVENTION: Recombinant HK2 Polypeptide
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner, P.A.
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: PCT/US95/06157
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Raasch, Kevin W.
; REGISTRATION NUMBER: 35,561
; REFERENCE/DOCKET NUMBER: 150.148W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-06157-1

Query Match 94.08; Score 47; DB 5; Length 237;
Best Local Similarity 80.08; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGGPXV 10
I I I I I I I I
Db 185 CSGDSGGPLV 194

RESULT 15

US-08-472-228A-1
; Sequence 1, Application US/08472228A
; Patent No. 5807978
; GENERAL INFORMATION:
; APPLICANT: Kokolus, William J.
; APPLICANT: Fritzsche, Herbert A.
; APPLICANT: Johnston, Dennis A.
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES OF PROSTATE
; TITLE OF INVENTION: SPECIFIC ANTIGEN
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,228A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSC:388
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-228A-1

Query Match 94.08; Score 47; DB 1; Length 240;
Best Local Similarity 80.08; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGGPXV 10
I I I I I I I I
Db 188 CSGDSGGPLV 197

Search completed: May 12, 2003, 15:40:37
Job time : 11.6 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:38:52 ; Search time 19.4 Seconds
(without alignments)
47.436 Million cell updates/sec

Title: US-09-909-348-2
Perfect score: 50
Sequence: 1 CXGDSGGPXV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	47	94.0	261	9	US-09-895-793-946
2	47	94.0	261	9	US-10-131-241-62
3	47	94.0	261	10	US-09-796-294-10
4	47	94.0	261	10	US-09-822-827-946
5	47	94.0	261	10	US-09-963-896-6
6	47	94.0	261	10	US-09-907-402-1
7	47	94.0	275	10	US-09-755-100-11
8	47	94.0	679	10	US-09-874-138-6
9	47	94.0	679	10	US-09-874-238-6
10	47	94.0	1019	9	US-10-183-992-4
11	47	94.0	1019	9	US-10-183-992-8
12	47	94.0	1079	9	US-09-895-793-947
13	47	94.0	1079	10	US-09-822-827-947
14	47	94.0	1083	9	US-10-183-992-6
15	46	92.0	10	9	US-10-050-692-1
16	46	92.0	10	9	US-10-050-692-2
17	46	92.0	10	9	US-10-050-688-1
18	46	92.0	10	9	US-10-050-688-2
19	46	92.0	10	10	US-09-909-348-1

20	46	92.0	10	10	US-09-909-348-2	Sequence 2, Appli
21	46	92.0	10	10	US-09-909-122-1	Sequence 1, Appli
22	46	92.0	10	10	US-09-909-122-2	Sequence 2, Appli
23	46	92.0	12	9	US-10-050-611-2	Sequence 2, Appli
24	46	92.0	12	9	US-10-125-459-18	Sequence 18, Appli
25	46	92.0	12	9	US-10-125-459-19	Sequence 19, Appli
26	46	92.0	12	9	US-10-067-761-33	Sequence 33, Appli
27	46	92.0	12	9	US-10-067-761-34	Sequence 34, Appli
28	46	92.0	12	10	US-09-879-792-16	Sequence 16, Appli
29	46	92.0	12	10	US-09-904-090-2	Sequence 2, Appli
30	46	92.0	12	10	US-09-904-090-4	Sequence 4, Appli
31	46	92.0	12	10	US-09-804-156-33	Sequence 33, Appli
32	46	92.0	12	10	US-09-804-156-34	Sequence 34, Appli
33	46	92.0	12	10	US-09-946-633-18	Sequence 18, Appli
34	46	92.0	12	10	US-09-946-633-19	Sequence 19, Appli
35	46	92.0	14	9	US-10-050-692-4	Sequence 4, Appli
36	46	92.0	14	9	US-10-050-688-4	Sequence 4, Appli
37	46	92.0	14	10	US-09-909-122-4	Sequence 4, Appli
38	46	92.0	14	10	US-09-885-441-6	Sequence 6, Appli
39	46	92.0	18	9	US-09-826-290-247	Sequence 247, App
40	46	92.0	18	9	US-10-045-170A-16	Sequence 16, Appli
41	46	92.0	18	9	US-10-044-708A-29	Sequence 29, Appli
42	46	92.0	23	9	US-10-050-692-6	Sequence 6, Appli
43	46	92.0	23	9	US-10-050-611-3	Sequence 3, Appli
44	46	92.0	23	9	US-10-050-611-4	Sequence 4, Appli
45	46	92.0	23	9	US-10-050-688-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-895-793-946
; Sequence 946, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-946
Query Match 94.0%; Score 47; DB 9; Length 261;
Best Local Similarity 80.0%; Pred. No. 1.3;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10
| | | | | | | |
Db 209 CSGDSGGPLV 218

RESULT 2

US-10-131-241-62
; Sequence 62, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-62

Query Match 94.0%; Score 47; DB 9; Length 261;
Best Local Similarity 80.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10
| | | | | | | |
Db 209 CSGDSGGPLV 218

RESULT 3

US-09-796-294-10
; Sequence 10, Application US/09796294
; Patent No. US20020037581A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: Extracellular Serine Protease
; FILE REFERENCE: D6020CIP3
; CURRENT APPLICATION NUMBER: US/09/796,294
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/618,259
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 10
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of Prostate Specific Antigen
; OTHER INFORMATION: (hpsa); accession no. P07288
US-09-796-294-10

Query Match 94.0%; Score 47; DB 10; Length 261;
Best Local Similarity 80.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10
| | | | | | | |
Db 209 CSGDSGGPLV 218

RESULT 4

US-09-822-827-946
; Sequence 946, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jlangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-946

Query Match 94.0%; Score 47; DB 10; Length 261;
Best Local Similarity 80.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGDSGGPXV 10
| | | | | | | |
Db 209 CSGDSGGPLV 218

RESULT 5

US-09-963-896-6
; Sequence 6, Application US/09963896
; Patent No. US20020102585A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/963,896
; FILING DATE: 26-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/397,558
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: CERONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0527 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 261 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank

US-09-963-896-6
; Sequence 6, Application US/09963896
; Patent No. US20020102585A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/963,896
; FILING DATE: 26-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/397,558
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: CERONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0527 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 261 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank


```

RESULT 10
US-10-183-992-4
; Sequence 4, Application US/10183992
; Publication No. US20030054432A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Lin
; APPLICANT: Pepe, Michael
; TITLE OF INVENTION: Methods and Reagents for Detecting Endotoxin
; FILE REFERENCE: 03877.00008
; CURRENT APPLICATION NUMBER: US/10/183.992
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 60/310,125
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1019
; TYPE: PRT
; ORGANISM: Tachypleudus tridentata
US-10-183-992-4
Query Match          94.0%; Score 47; DB 9; Length 1019;
Best Local Similarity 80.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXGDSGGPXV 10
Db 962 CSGDSGGPLV 971

```

```

RESULT 11
US-10-183-992-8
; Sequence 8, Application US/10183992
; Publication No. US20030054432A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Lin
; APPLICANT: Pepe, Michael
; TITLE OF INVENTION: Methods and Reagents for Detecting Endotoxin
; FILE REFERENCE: 03877.00008
; CURRENT APPLICATION NUMBER: US/10/183.992
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 60/310,125
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1019
; TYPE: PRT
; ORGANISM: Carcinoscopus rotundicauda
US-10-183-992-8
Query Match          94.0%; Score 47; DB 9; Length 1019;
Best Local Similarity 80.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXGDSGGPXV 10
Db 962 CSGDSGGPLV 971

```

```

RESULT 12
US-09-895-793-947
; Sequence 947, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.

```

```

; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Repler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895.793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 947
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-947
Query Match          94.0%; Score 47; DB 9; Length 1079;
Best Local Similarity 80.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXGDSGGPXV 10
Db 216 CSGDSGGPLV 225

```

```

RESULT 13
US-09-822-827-947
; Sequence 947, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822.827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 947
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-947
Query Match          94.0%; Score 47; DB 10; Length 1079;
Best Local Similarity 80.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXGDSGGPXV 10
Db 216 CSGDSGGPLV 225

```

```

RESULT 14
US-10-183-992-6
; Sequence 6, Application US/10183992
; Publication No. US20030054432A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Lin
; APPLICANT: Pepe, Michael
; TITLE OF INVENTION: Methods and Reagents for Detecting Endotoxin

```

; FILE REFERENCE: 02877.00008
; CURRENT APPLICATION NUMBER: US/10/183,992
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 60/310,125
; PRIOR FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1083
; TYPE: PRT
; ORGANISM: Carinoscorpius rotundicauda
US-10-183-992-6

Query Match 94.0%; Score 47; DB 9; Length 1083;
Best Local Similarity 80.0%; Pred. NO. 5.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGP XV 10
| | | | | | | | | |
Db 1026 CSGDSGGP LV 1035

RESULT 15

US-10-050-692-1
; Sequence 1, Application US/10050692
; Publication No. US20020182205A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Simmons, David J.
; APPLICANT: Yang, Jiping
; APPLICANT: Redin, William R.
; TITLE OF INVENTION: STIMULATION OF BONE GROWTH WITH THROMBIN
; TITLE OF INVENTION: PEPTIDE DERIVATIVES
; FILE REFERENCE: 3033.1002-004
; CURRENT APPLICATION NUMBER: US/10/050,692
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 09/909,122
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/219,300
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human prothrombin
US-10-050-692-1

Query Match 92.0%; Score 46; DB 9; Length 10;
Best Local Similarity 80.0%; Pred. NO. 0.089;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGP XV 10
| | | | | | | | | |
Db 1 CEGDSGGP V 10.

Search completed: May 12, 2003, 15:59:09
Job time : 20.4 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:33:02 ; Search time 11.4 Seconds
(without alignments)
84.328 Million cell updates/sec

Title: US-09-909-348-2

Perfect score: 50

Sequence: 1 CXGDSGGPXV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	94.0	251	2	29K serine protein
2	47	94.0	254	2	chymotrypsin-like
3	47	94.0	261	1	semenogelase (EC 3
4	47	94.0	261	1	semenogelase (EC 3
5	47	94.0	366	2	testicular serine
6	47	94.0	699	1	Ra-reactive factor
7	47	94.0	1019	2	serine proteinase
8	46	92.0	61	2	callicrein - mouse
9	46	92.0	66	2	elastase (EC 3.4.2
10	46	92.0	73	2	chymotrypsin-like
11	46	92.0	126	2	tissue kallikrein
12	46	92.0	149	1	coagulation factor
13	46	92.0	161	2	complement factor
14	46	92.0	162	2	complement factor
15	46	92.0	191	2	chymotrypsin (EC 3
16	46	92.0	216	1	chymotrypsin (EC 3
17	46	92.0	218	1	chymotrypsin (EC 3
18	46	92.0	225	2	probable serine pr
19	46	92.0	229	1	trypsin (EC 3.4.21
20	46	92.0	229	1	trypsin (EC 3.4.21
21	46	92.0	230	2	hypodermin C (EC 3
22	46	92.0	231	1	trypsin (EC 3.4.21
23	46	92.0	231	2	trypsin (EC 3.4.21
24	46	92.0	234	2	thrombin (EC 3.4.2
25	46	92.0	235	2	thrombin (EC 3.4.2
26	46	92.0	235	2	thrombin (EC 3.4.2
27	46	92.0	235	2	thrombin (EC 3.4.2
28	46	92.0	236	2	thrombin (EC 3.4.2
29	46	92.0	236	2	thrombin (EC 3.4.2

30	46	92.0	237	2	S68702	trypsin (EC 3.4.2
31	46	92.0	238	1	TRW5Y	trypsin-like prote
32	46	92.0	238	2	S31779	trypsin (EC 3.4.21
33	46	92.0	239	2	G42696	thrombin (EC 3.4.2
34	46	92.0	240	2	S39047	trypsin (EC 3.4.21
35	46	92.0	241	2	S39048	trypsin (EC 3.4.21
36	46	92.0	242	2	S31775	trypsin (EC 3.4.21
37	46	92.0	242	2	S49489	trypsin (EC 3.4.21
38	46	92.0	242	2	S31776	trypsin (EC 3.4.21
39	46	92.0	243	2	A35871	trypsin (EC 3.4.21
40	46	92.0	244	2	S72219	chymotrypsin B - A
41	46	92.0	245	1	KYBOB	chymotrypsin (EC 3
42	46	92.0	245	1	KYBOA	chymotrypsin (EC 3
43	46	92.0	245	2	A48598	kallikrein-like se
44	46	92.0	246	1	TRRT1	trypsin (EC 3.4.21
45	46	92.0	246	1	TRRT2	trypsin (EC 3.4.21

ALIGNMENTS

RESULT 1

PC1235
29K serine proteinase (EC 3.4.21.-) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-May-1997
C:Accession: PC1235; PC1236
R:Takahashi, A.; Takayama, Y.; Hatsuse, H.; Kawakami, M.
Biochem. Biophys. Res. Commun. 190, 681-687, 1993
A:Title: Presence of a serine protease in the complement-activating component of the
A:Reference number: PC1235; MUID:93176166; PMID:8493319
A:Accession: PC1235
A:Molecule type: mRNA
A:Residues: 1-251 <TAK>
A:Accession: PC1236
A:Molecule type: protein
A:Residues: 1-18,122-135,155-173;182-201 <TAK>
C:Superfamily: complement subcomponent C1r; C1r/C1s repeat homology; complement facto
C:Keywords: hydrolase; serine proteinase
F:1-243/Domain: trypsin homology <TRY>

Query Match 94.0%; Score 47; DB 2; Length 251;
Best Local Similarity 80.0%; Pred. No. 0.36; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 2;

Qy 1 CXGDSGGPXV 10
| | | | | | | |
Db 194 CAGDSGGPFW 203

RESULT 2

S35585
chymotrypsin-like proteinase (EC 3.4.21.-) precursor - California red abalone
C:Species: Haliotis rufescens (California red abalone)
C:Date: 07-Apr-1994 #sequence_revision 19-May-1994 #text_change 22-Jun-1999
C:Accession: S35585; S32750
R:Groppe, J.C.; Morse, D.E.
Arch. Biochem. Biophys. 305, 159-169, 1993
A:Title: Molluscan chymotrypsin-like protease: structure, localization, and substrate
A:Reference number: S35585; MUID:93343624; PMID:8342947
A:Accession: S35585
A:Molecule type: mRNA
A:Residues: 1-254 <GRO>
A:Cross-references: EMBL:X71438; NID:G296361; PIDN:CAA50572.1; PID:G296362
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; intestine; serine proteinase; zymogen
F:1-23/Domain: signal sequence #status predicted <SIG>
F:23-254/Product: chymotrypsin-like proteinase #status predicted <MAT>
F:24-251/Domain: trypsin homology <TRY>

Query Match 94.0%; Score 47; DB 2; Length 254;
Best Local Similarity 80.0%; Pred. No. 0.36; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 2;

QY 1 CXGSGGXPV 10
| | | | | | | |
Db 208 CSGSGGPLV 217

RESULT 3

A32297 semenogelase (EC 3.4.21.77) precursor [validated] - human
N;Alternate names: gamma-seminoprotein; p-30 antigen; prostate-specific antigen; prostatic
C;Species: Homo sapiens (man)
C;Date: 19-May-2000 #sequence.revision 19-May-2000 #text.change 08-Dec-2000
C;Accession: A32297; S03604; S05468; S05467; A32546; S02239; A26757; C31567; A31567
R;Riegman, P.H.J.; Vilesttra, R.J.; van der Korput, J.A.G.M.; Romijn, J.C.; Trapman, J.
Biochem. Biophys. Res. Commun. 159, 95-102, 1989
A;Title: Characterization of the prostate-specific antigen gene: a novel human kallikrein
A;Reference number: A32297; MUID:89165891; PMID:2466464
A;Accession: A32297
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-261 <RIE>
A;Cross-references: GB:M24543
R;Lundwall, A.
Biochem. Biophys. Res. Commun. 161, 1151-1159, 1989
A;Title: Characterization of the gene for prostate-specific antigen, a human glandular k
A;Reference number: A32423; MUID:89302090; PMID:2472789
A;Accession: A32423
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-261 <LUN>
A;Cross-references: GB:M27274; NID:9190552; PIDN:AAA60192.1; PID:9190553
R;Digby, M.; Zhang, X.Y.; Richards, R.I.
Nucleic Acids Res. 17, 2137, 1989
A;Title: Human prostate specific antigen (PSA) gene: structure and linkage to the kallik
A;Reference number: S03604; MUID:89183632; PMID:2467258
A;Accession: S03604
A;Molecule type: DNA
A;Residues: 1-261 <DIG>
A;Cross-references: EMBL:X13940
R;Klobeck, H.G.; Combrato, G.; Schulz, P.; Arbusow, V.; Fittler, F.
submitted to the EMBL Data Library, May 1989
A;Reference number: S05468
A;Accession: S05468
A;Molecule type: DNA
A;Residues: 1-261 <KLJ>
A;Cross-references: EMBL:X14810; NID:935732; PIDN:CAA32915.1; PID:g296671
R;Klobeck, H.G.; Combrato, G.; Schulz, P.; Arbusow, V.; Fittler, F.
Nucleic Acids Res. 17, 3981, 1989
A;Title: Genomic sequence of human prostate specific antigen (PSA).
A;Reference number: S05467; MUID:89282407; PMID:2471958
A;Accession: S05467
A;Molecule type: DNA
A;Residues: 1-29 <KLJ>
A;Cross-references: EMBL:X14810
R;Henttu, P.; Viikko, P.
Biochem. Biophys. Res. Commun. 160, 903-910, 1989
A;Title: cDNA coding for the entire human prostate specific antigen shows high homologie
A;Reference number: A32546; MUID:89246551; PMID:2470373
A;Accession: A32546
A;Molecule type: mRNA
A;Residues: 1-72, 'T', 74-85, 'I', 87-174, 'P', 176-183, 'Q', 185-259, 'D', 261 <HEN>
A;Cross-references: GB:M26663
R;Schulz, P.; Stucka, R.; Feldmann, H.; Combrato, G.; Klobeck, H.G.; Fittler, F.
Nucleic Acids Res. 16, 6226, 1988
A;Title: Sequence of a cDNA clone encompassing the complete mature human prostate specif
A;Reference number: S02239; MUID:88289366; PMID:2456523
A;Accession: S02239
A;Molecule type: mRNA
A;Residues: 17-63, 'T', 65-135, 'W', 137-261 <SCH>
A;Cross-references: EMBL:X07730
R;Lundwall, A.; Lilja, H.
FEBS Lett. 214, 317-322, 1987
A;Title: Molecular cloning of human prostate specific antigen cDNA.

A;Reference number: A26757; MUID:87190978; PMID:2436946
A;Accession: A26757
A;Molecule type: mRNA
A;Residues: 5-261 <LU2>
R;Riegman, P.H.J.; Klaassen, P.; van der Korput, J.A.G.M.; Romijn, J.C.; Trapman, J.
Biochem. Biophys. Res. Commun. 155, 181-188, 1988
A;Title: Molecular cloning and characterization of novel prostate antigen cDNA's.
A;Reference number: A90144; MUID:88326297; PMID:2458104
A;Accession: C31567
A;Molecule type: mRNA
A;Residues: 5-261 <RI2>
A;Cross-references: GB:M21895; NID:9189523; PIDN:AAA59995.1; PID:g189524
A;Note: the authors translated the codon GGC for residue 28 as Arg and TGG for residu
A;Accession: A31567
A;Molecule type: mRNA
A;Residues: 5-210, 'VVILITELTTPA', 223, 'PMVLHGLV', 233, 'WRGGV' <RI3>
A;Cross-references: GB:M21896; NID:9189525; PIDN:AAA59996.1; PID:g189526
A;Note: the authors translated the codon GGC for residue 28 as Arg and TGG for residu
A;Accession: B31567
A;Molecule type: mRNA
A;Residues: 5-164, 'CTPGDGAAGSPDAWV' <RI4>
A;Cross-references: GB:M21897; NID:9189529; PIDN:AAA59997.1; PID:g189530
A;Note: the authors translated the codon GGC for residue 28 as Arg, TGG for residue 2
R;Schaller, J.; Akiyama, K.; Tsuda, R.; Hara, M.; Marti, T.; Rickli, E.E.
Eur. J. Biochem. 170, 111-120, 1987
A;Title: Isolation, characterization and amino-acid sequence of gamma-seminoprotein,
A;Reference number: S00232; MUID:88082806; PMID:3691515
A;Accession: S00232
A;Molecule type: protein
A;Residues: 25-261 <SC2>
R;Watt, K.W.K.; Lee, P.J.; M'Timkulu, T.; Chan, W.P.; Looor, R.
Proc. Natl. Acad. Sci. U.S.A. 83, 3166-3170, 1986
A;Title: Human prostate-specific antigen: structural and functional similarity with s
A;Reference number: A23937; MUID:86205857; PMID:2422647
A;Accession: A23937
A;Molecule type: protein
A;Residues: 25-93, 'T', 95-164, 'HL', 166, 'YDQM', 169-174, 'Q', 176-261 <WAT>
R;Moreno, J.M.
submitted to the EMBL Data Library, November 1994
A;Reference number: G07735
A;Accession: G01551
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-261 <WOR>
A;Cross-references: EMBL:U17040; NID:9595945; PIDN:AAA56764.1; PID:g595946
R;Monne, M.; Croce, C.M.; Yu, H.; Diamandis, E.P.
Cancer Res. 54, 6344-6347, 1994
A;Title: Molecular characterization of prostate-specific antigen messenger RNA expres
A;Reference number: I52712; MUID:95079406; PMID:7527295
A;Accession: I52712
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'POAYLHPSCVT', 1-176 <RES>
A;Cross-references: GB:S75755; NID:9861469; PIDN:AAD14185.1; PID:g4261885
R;Christenson, A.; Lilja, H.
Eur. J. Biochem. 220, 45-53, 1994
A;Title: Complex formation between protein C inhibitor and prostate-specific antigen
A;Reference number: S41212; MUID:94164172; PMID:7509746
A;Accession: S41212
A;Molecule type: protein
A;Residues: 25-30, 'X', 32-49 <CHR>
A;Comment: This enzyme preferentially cleaves after tyrosine residues.
C;Genetics:
A;Gene: GDB:APS; PSA
A;Cross-references: GDB:119695; OMIM:176820
A;Map position: 19q13.3-19q13.3
A;Introns: 16/1; 69/2; 165/1; 210/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; prostate; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-24/Domain: propeptide #status predicted <PRO>
F;25-261/Product: semenogelase #status experimental <WAT>
F;25-253/Domain: trypsin homology <TRY>

F:65,120,213/Active site: His, Asp, Ser #status predicted

Query Match 94.0%; Score 47; DB 1; Length 261;
Best Local Similarity 80.0%; Pred. No. 0.36;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
| | | | | | | | | |
DB 209 CSGDSGGPLV 218

RESULT 4

S35711 semenogelase (EC 3.4.21.77) precursor - rhesus macaque

N:Alternate names: gamma-seminoprotein; P-30 antigen; prostate-specific antigen; prostatic

C:Species: Macaca mulatta (rhesus macaque)

C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-May-2000

C:Accession: S35711; S34239

R:Gauthier, E.R.; Chapdelaine, P.; Tremblay, R.R.; Dube, J.Y.

Biochim. Biophys. Acta 1174, 207-210, 1993

A:Title: Characterization of rhesus monkey prostate specific antigen cDNA.

A:Reference number: S35711; MUID:93363642; PMID:7689340

A:Accession: S35711

A:Molecule type: mRNA

A:Residues: 1-261 <GAD>

A:Cross-references: EMBL:X73560; NID:g311843; PIDN:CAA51957.1; PID:g311844

C:Comment: This enzyme preferentially cleaves after tyrosine residues.

C:Superfamily: trypsin; trypsin homology

C:Keywords: glycoprotein; hydrolase; prostate; serine proteinase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-24/Domain: propeptide #status predicted <PRO>

F:25-261/Product: semenogelase #status predicted <MAT>

F:25-253/Domain: trypsin homology <TRY>

F:65,120,213/Active site: His, Asp, Ser #status predicted

Query Match 94.0%; Score 47; DB 1; Length 261;
Best Local Similarity 80.0%; Pred. No. 0.36;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
| | | | | | | | | |
DB 209 CSGDSGGPLV 218

RESULT 5

JE0105

testicular serine proteinase (EC 3.4.21.-) TESP2 - mouse

N:Alternate names: TESP2

C:Species: Mus musculus (house mouse)

C:Date: 22-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000

C:Accession: JE0105

R:Kohn, N.; Yamagata, K.; Yamada, S.; Kashiwabara, S.; Sakai, Y.; Baba, T.

Biochem. Biophys. Res. Commun. 245, 658-665, 1998

A:Title: Two novel testicular serine proteinases, TESP1 and TESP2, are present in the mouse

A:Reference number: JE0104; MUID:98249770; PMID:9588171

A:Accession: JE0105

A:Molecule type: mRNA

A:Residues: 1-366 <KOH>

A:Cross-references: DBJ:AB008911; NID:g3130174; PIDN:BA326133.1; PID:g3130175

C:Comment: This protein is localized in the sperm acrosome and is released during the ac

C:Superfamily: trypsin homology

C:Keywords: hydrolase; serine proteinase

F:71-310/Domain: trypsin homology <TRY>

Query Match 94.0%; Score 47; DB 2; Length 366;
Best Local Similarity 80.0%; Pred. No. 0.5;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
| | | | | | | | | |
DB 262 CAGDSGGPLV 271

RESULT 6

I54763

Ra-reactive factor (EC 3.4.21.-) 1 precursor - human

N:Alternate names: mannose binding protein-associated serine proteinase 1 (MASP-1)

C:Species: Homo sapiens (man)

C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 16-Jun-2000

C:Accession: I54763; JN0883

R:Sato, T.; Endo, Y.; Matsushita, M.; Fujita, T.

Int. Immunol. 6, 665-669, 1994

A:Title: Molecular characterization of a novel serine protease involved in activation

A:Reference number: I54763; MUID:94289349; PMID:8018603

A:Accession: I54763

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-699 <SAT>

A:Cross-references: GB:D28593; NID:g790963; PIDN:BAA05928.1; PID:g471128

R:Takada, F.; Takayama, Y.; Hattuse, H.; Kawakami, M.

Biochem. Biophys. Res. Commun. 196, 1003-1009, 1993

A:Title: A new member of the C1s family of complement proteins found in a bactericida

A:Reference number: JN0883; MUID:94059062; PMID:8240317

A:Accession: JN0883

A:Molecule type: mRNA

A:Residues: 1-234, 'E', 236-284, 'G', 286-498, 'K', 500-542, 'K', 544-642, 'S', 644-699 <TAK>

A:Cross-references: DBJ:DI17525; NID:g439712; PIDN:BAA04477.1; PID:g439713

A:Experimental source: liver

C:Comment: This is a serum bactericidal factor that activates complement C4 and C2 co

C:Genetics:

A:Gene: GDB:MSP1; GDB:CRARF; CRARF1; PRSS5; MASP

A:Cross-references: GDB:361104; GDB:330954; OMIM:600521

A:Map position: 3q27-3q28

C:Superfamily: complement subcomponent C1r; C1r/C1s repeat homology; complement factor

C:Keywords: beta-hydroxyasparagine; complement pathway; duplication; glycoprotein; hy

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-448,449-699/Product: Ra-reactive factor #status predicted <MAT>

F:19-135/Domain: C1r/C1s repeat homology <C1r1>

F:143-181/Domain: EGF homology <EGF>

F:185-234/Domain: C1r/C1s repeat homology <C1r2>

F:301-362/Domain: complement factor H repeat homology <FHL>

F:367-432/Domain: complement factor H repeat homology <FH2>

F:449-691/Domain: trypsin homology <TRY>

F:49,178,407/Binding site: carbohydrtase (Asn) (covalent) #status predicted

F:73-91,143-157,153-166,168-181,185-212,242-260,301-349,329-362,367-414,397-432,436-5

F:159/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

F:448-449/cleavage site: Arg-Ile (autolytic) #status predicted

F:490,552,646/Active site: His, Asp, Ser #status predicted

Query Match 94.0%; Score 47; DB 1; Length 699;

Best Local Similarity 80.0%; Pred. No. 0.92;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
| | | | | | | | | |
DB 642 CAGDSGGPMV 651

RESULT 7

A38738

coagulation factor C precursor - horseshoe crab (Tachyples tridentatus)

N:Alternate names: coagulation-complement factor C; Limulus factor C

N:Contains: coagulation factor C heavy chain; coagulation factor C light chain peptid

C:Species: Tachyples tridentatus

C:Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 08-Dec-2000

C:Accession: A38738; B38738; S00105

R:Muta, T.; Miyata, T.; Misumi, Y.; Tokunaga, F.; Nakamura, T.; Toh, Y.; Ikehara, Y.;

J. Biol. Chem. 266, 6554-6561, 1991

A:Title: Limulus factor C. An endotoxin-sensitive serine protease zymogen with a mosa

A:Reference number: A38738; MUID:91177916; PMID:2007602

A:Accession: A38738

A:Molecule type: mRNA

A:Residues: 1-1019 <MUT>

A:Cross-references: GB:D90271; NID:g217396; PID:d1015020; PID:g217397

A:Accession: B38738

A:Molecule type: mRNA

A;Residues: 1-466,616,'DN',619-620,'A',622 <MU2>
 A;Cross-references: GB:D90272
 R;Tokunaga, F.; Miyata, T.; Nakamura, T.; Morita, T.; Kuma, K.I.; Miyata, T.; Iwanaga, S.
 Eur. J. Biochem. 167, 405-416, 1987
 A;Title: Lipopolysaccharide-sensitive serine-protease zymogen (factor C) of horseshoe crab
 is a novel type of serine protease.
 A;Reference number: S00105; MUID:88004461; PMID:3308457
 A;Accession: S00105
 A;Molecule type: protein
 A;Residues: 'S',27-39,'S',41,'T',43-54,'E',691-782;950-977 <TOK>
 C;Superfamily: coagulation factor C; C-type lectin homology; complement factor H repeat
 C;Keywords: alternative splicing; glycoprotein; hemolymph coagulation; hydrolase; serine
 F;1-25/Domain: signal sequence #status predicted <SIG>
 F;26-60/Product: coagulation factor C heavy chain #status experimental <HCH>
 F;136-195/Domain: complement factor H repeat homology <FH01>
 F;199-254/Domain: complement factor H repeat homology <FH02>
 F;260-321/Domain: complement factor H repeat homology <FH03>
 F;436-564/Domain: C-type lectin homology <LCH>
 F;576-634/Domain: complement factor H repeat homology <FH04>
 F;685-747/Domain: complement factor H repeat homology #status atypical <FH05>
 F;691-762/Product: coagulation factor C light chain peptide A #status experimental <PPA>
 F;763-1019/Product: coagulation factor C light chain peptide B #status experimental <PPB>
 F;763-1015/Domain: trypsin homology <TRY>
 F;523,534,624,912/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;767/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F;809,865,966/Active site: His, Asp, Ser #status predicted

Query Match 94.0%; Score 47; DB 2; Length 1019;
 Best Local Similarity 80.0%; Pred. No. 1.3;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | ||||| |
 Db 962 CSGDSGGPLV 971

RESULT 8
 PS0049
 serine proteinase (EC 3.4.21.-) 3 - fruit fly (*Drosophila melanogaster*) (fragment)
 C;Species: *Drosophila melanogaster*
 C;Date: 07-Jun-1990 #sequence_revision 31-Dec-1991 #text_change 22-Jun-1999
 C;Accession: PS0049
 R;Yun, Y.; Davis, R.L.
 Mol. Cell. Biol. 9, 692-700, 1989
 A;Title: Levels of RNA from a family of putative serine protease genes are reduced in *D.*
 A;Reference number: JS0260; MUID:89219063; PMID:2469005
 A;Accession: PS0049
 A;Molecule type: DNA
 A;Residues: 1-61 <YUN>
 A;Cross-references: GB:M24380; NID:g158409; PIDN:RAA02551.1; PID:g158412
 A;Experimental source: strain dnc mutant
 C;Genetics:
 A;Gene: SER3
 A;Cross-references: FlyBase:FBgn0003358
 A;Map position: 3 99C-D
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; serine proteinase
 F;1-53/Domain: trypsin homology (fragment) <TRY>
 F;30/Binding site: substrate (Val) #status predicted

Query Match 92.0%; Score 46; DB 2; Length 61;
 Best Local Similarity 80.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | ||||| |
 Db 7 CQGDGGPLV 16

RESULT 9
 I52972
 kallikrein - mouse (fragment)
 C;Species: *Mus musculus* (house mouse)

A;Residues: 1-466,616,'DN',619-620,'A',622 <MU2>
 A;Cross-references: GB:D90272
 R;Tokunaga, F.; Miyata, T.; Nakamura, T.; Morita, T.; Kuma, K.I.; Miyata, T.; Iwanaga, S.
 Eur. J. Biochem. 167, 405-416, 1987
 A;Title: Lipopolysaccharide-sensitive serine-protease zymogen (factor C) of horseshoe crab
 is a novel type of serine protease.
 A;Reference number: S00105; MUID:88004461; PMID:3308457
 A;Accession: S00105
 A;Molecule type: protein
 A;Residues: 'S',27-39,'S',41,'T',43-54,'E',691-782;950-977 <TOK>
 C;Superfamily: coagulation factor C; C-type lectin homology; complement factor H repeat
 C;Keywords: alternative splicing; glycoprotein; hemolymph coagulation; hydrolase; serine
 F;1-25/Domain: signal sequence #status predicted <SIG>
 F;26-60/Product: coagulation factor C heavy chain #status experimental <HCH>
 F;136-195/Domain: complement factor H repeat homology <FH01>
 F;199-254/Domain: complement factor H repeat homology <FH02>
 F;260-321/Domain: complement factor H repeat homology <FH03>
 F;436-564/Domain: C-type lectin homology <LCH>
 F;576-634/Domain: complement factor H repeat homology <FH04>
 F;685-747/Domain: complement factor H repeat homology #status atypical <FH05>
 F;691-762/Product: coagulation factor C light chain peptide A #status experimental <PPA>
 F;763-1019/Product: coagulation factor C light chain peptide B #status experimental <PPB>
 F;763-1015/Domain: trypsin homology <TRY>
 F;523,534,624,912/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;767/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F;809,865,966/Active site: His, Asp, Ser #status predicted

C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 22-Jun-1999
 C;Accession: I52972
 R;Norden, S.K.; Mason, A.J.; Richards, R.I.; Baxter, J.D.; Shine, J.
 DNA 1, 309-311, 1982
 A;Title: Mouse kallikrein arginyl-esteropeptidase genes: Analysis of cloned cDNAs sug
 A;Reference number: I52972; MUID:83182015; PMID:6926406
 A;Accession: I52972
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-66 <RES>
 A;Cross-references: GB:K01654; NID:g198490; PIDN:AAA39346.1; PID:g198491
 C;Superfamily: trypsin; trypsin homology
 F;1-58/Domain: trypsin homology (fragment) <TRY>

Query Match 92.0%; Score 46; DB 2; Length 66;
 Best Local Similarity 70.0%; Pred. No. 0.15;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | ||||| |
 Db 14 CAGDSGGPLI 23

RESULT 10
 S44462
 elastase (EC 3.4.21.-) 2B - horse (fragments)
 C;Species: *Equus caballus* (domestic horse)
 C;Date: 20-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 07-May-1999
 C;Accession: S44462
 R;Dubin, A.; Potempa, J.; Travis, J.
 Biochem. J. 300, 401-406, 1994
 A;Title: Structural and functional characterization of elastases from horse neutrophil
 A;Reference number: S44461; MUID:94271153; PMID:7516152
 A;Accession: S44462
 A;Molecule type: protein
 A;Residues: 1-31;32-56;57-73 <DUB>
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; serine proteinase
 F;1-73/Domain: trypsin homology (fragments) <TRY>

Query Match 92.0%; Score 46; DB 2; Length 73;
 Best Local Similarity 80.0%; Pred. No. 0.17;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | ||||| |
 Db 60 CFGDSGGPLV 69

RESULT 11
 A23473
 chymotrypsin-like proteinase (EC 3.4.21.-) - pig (tentative sequence) (fragments)
 N;Alternate names: pancreatic elastase II [misidentification]
 C;Species: *Sus scrofa domestica* (domestic pig)
 C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 31-Mar-2000
 C;Accession: A23473
 R;Vered, M.; Gertler, A.; Burstein, Y.
 Int. J. Pept. Protein Res. 27, 183-190, 1986
 A;Reference number: A23473; MUID:86194934; PMID:3634756
 A;Accession: A23473
 A;Molecule type: protein
 A;Residues: 1-126 <VER>
 C;Superfamily: trypsin; trypsin homology
 C;Keywords: hydrolase; serine proteinase

Query Match 92.0%; Score 46; DB 2; Length 126;
 Best Local Similarity 80.0%; Pred. No. 0.28;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | ||||| |
 Db 96 CQGDGGPLV 105

RESULT 12

KQMSM
tissue kallikrein (EC 3.4.21.35); submandibular - mouse (fragment)
N:Alternate names: glandular kallikrein; kininogenin
C:Species: Mus musculus (house mouse)
C:Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 16-Jun-2000
C:Accession: A00939
R:Richards, R.I.; Catanzaro, D.F.; Mason, A.J.; Morris, B.J.; Baxter, J.D.; Shine, J.
J. Biol. Chem. 257, 2758-2761, 1982
A:Title: Mouse glandular kallikrein genes. Nucleotide sequence of cloned cDNA coding for
A:Reference number: A00939; MUID:82142394; PMID:6174512
A:Accession: A00939
A:Molecule type: mRNA
A:Residues: 1-149 <RIC>
A:Cross-references: GB:V00828; GB:J00389; NID:g52773; PIDN:CAA24211.1; PID:g1334112
A:Experimental source: Quakenbush inbred strain
C:Comment: Tissue kallikreins cleave Met-Lys and Arg-Ser bonds in kininogen to release
C:Genetics:
A:Map position: 7
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; saliva; serine proteinase; submandibular gland
F:1-141/Domain: trypsin homology (fragment) <TRY>
F:8,101/Active site: Asp, Ser #status predicted
F:40-107,72-86,97-122/Disulfide bonds: #status predicted

Query Match 92.0%; Score 46; DB 1; Length 149;
Best Local Similarity 70.0%; Pred. No. 0.33;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
I | | | | | | |
DB 97 CAGDSGGPLI 106

RESULT 13

I62744
coagulation factor Xa (EC 3.4.21.6) - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 22-Jun-1999
C:Accession: I62744
R:Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.
Eur. J. Haematol. 52, 162-168, 1994
A:Title: Analysis of the partial nucleotide sequences and deduced primary structures of
A:Reference number: I46196; MUID:94222160; PMID:8168596
A:Accession: I62744
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-161 <RES>
A:Cross-references: GB:D21214; NID:g415307; PIDN:BAA04755.1; PID:g455395
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutamic
F:1-161/Domain: trypsin homology (fragment) <TRY>

Query Match 92.0%; Score 46; DB 2; Length 161;
Best Local Similarity 80.0%; Pred. No. 0.35;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
I | | | | | | |
DB 119 CQGDGGPHV 128

RESULT 14

I48158
coagulation factor Xa (EC 3.4.21.6) - golden hamster (fragment)
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1999
C:Accession: I48158
R:Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.
Eur. J. Haematol. 52, 162-168, 1994
A:Title: Analysis of the partial nucleotide sequences and deduced primary structures of
A:Reference number: I46196; MUID:94222160; PMID:8168596

A:Accession: I48158
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-161 <RES>
A:Cross-references: GB:D21216; NID:g415304; PIDN:BAA04757.1; PID:g455393
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu
F:1-161/Domain: trypsin homology (fragment) <TRY>

Query Match 92.0%; Score 46; DB 2; Length 161;
Best Local Similarity 80.0%; Pred. No. 0.35;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
I | | | | | | |
DB 119 CQGDGGPHV 128

RESULT 15

S54115
complement factor D (EC 3.4.21.46) - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 19-May-2000
C:Accession: S54115
R:Nicolas, N.
submitted to the EMBL Data Library, April 1995
A:Reference number: S54115
A:Accession: S54115
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-191 <NIC>
A:Cross-references: EMBL:249058; NID:g773264; PIDN:CAA88844.1; PID:g773265
C:Superfamily: trypsin; trypsin homology
C:Keywords: complement alternate pathway; hydrolase; plasma; serine proteinase
F:1-181/Domain: trypsin homology (fragment) <TRY>

Query Match 92.0%; Score 46; DB 2; Length 191;
Best Local Similarity 80.0%; Pred. No. 0.41;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
I | | | | | | |
DB 137 CKGDGGPLV 146

Search completed: May 12, 2003, 15:39:40
Job time : 11.4 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:32:07 ; Search time 5.8 Seconds
(without alignments)
71.511 Million cell updates/sec

Title: US-09-909-348-2
Perfect score: 50
Sequence: 1 CXGDSGGPXV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	94.0	252	1 TRY1_DROME	P52905 drosophila
2	47	94.0	254	1 CTRB_HUMAN	P35003 haliotis ru
3	47	94.0	261	1 KLK3_HUMAN	P07288 homo sapien
4	47	94.0	261	1 KLK3_MACMU	P35619 macaca mula
5	47	94.0	699	1 CRAR_HUMAN	P48740 h complen
6	47	94.0	704	1 CRAR_MOUSE	P98064 mus musculu
7	47	94.0	1019	1 LFC_CARRO	Q26422 carcinoscor
8	47	94.0	1019	1 LFC_TACTR	P28175 tachyplesu
9	46	92.0	73	1 EL2B_HORSE	P37358 equus cabal
10	46	92.0	157	1 PRTC_CANFA	Q28278 canis famil
11	46	92.0	157	1 PRTC_CAPHI	Q28315 capra hircu
12	46	92.0	157	1 PRTC_FELCA	Q28412 felis silve
13	46	92.0	157	1 PRTC_HORSE	Q28380 equus cabal
14	46	92.0	161	1 PRTC_MACMU	Q28506 macaca mula
15	46	92.0	165	1 TRY3_LUCCU	P35043 lucilia cup
16	46	92.0	216	1 CTR2_VESOR	P00768 vespa orien
17	46	92.0	218	1 CTR2_VESCR	P00769 vespa crabr
18	46	92.0	227	1 TRIP_SACER	P24664 saccharopol
19	46	92.0	229	1 TRYP_SQUAC	P00764 squalus aca
20	46	92.0	231	1 TRYP_SALSA	P35032 salmo salar
21	46	92.0	231	1 TRYP_PIG	P00761 sus scrofa
22	46	92.0	235	1 TRYD_HUMAN	Q9bz13 homo sapien
23	46	92.0	238	1 TRY3_SALSA	P35033 salmo salar
24	46	92.0	238	1 TRY5_AEDAE	P29787 aedes aegypt
25	46	92.0	241	1 TRY1_GADMO	P16049 gadus morhu
26	46	92.0	241	1 TRYX_GADMO	Q91041 gadus morhu
27	46	92.0	242	1 TRY1_SALSA	P35031 salmo salar
28	46	92.0	243	1 TRY1_BOVIN	P00760 bos taurus
29	46	92.0	243	1 TRY1_XENLA	P19799 xenopus lae
30	46	92.0	244	1 KLK6_HUMAN	Q92876 homo sapien
31	46	92.0	244	1 TRY2_XENLA	P70059 xenopus lae
32	46	92.0	245	1 CTRA_BOVIN	P00766 bos taurus
33	46	92.0	245	1 CTRB_BOVIN	P00767 bos taurus

34	46	92.0	245	1 CTRB_GADMO	P80646 gadus morhu
35	46	92.0	245	1 GILX_HELHO	P43685 heloderma h
36	46	92.0	246	1 TRY1_CANFA	P06871 canis famil
37	46	92.0	246	1 TRY1_RAT	P00762 rattus norv
38	46	92.0	246	1 TRY2_MOUSE	P07146 mus musculu
39	46	92.0	246	1 TRY2_RAT	P00763 rattus norv
40	46	92.0	247	1 TRY1_HUMAN	P07477 homo sapien
41	46	92.0	247	1 TRY2_CANFA	P06872 canis famil
42	46	92.0	247	1 TRY2_HUMAN	P07478 homo sapien
43	46	92.0	247	1 TRY3_RAT	P08426 rattus norv
44	46	92.0	247	1 TRY4_RAT	P12788 rattus norv
45	46	92.0	248	1 KLKC_HUMAN	Q9ukr0 homo sapien

ALIGNMENTS

RESULT 1

TRY1_DROME
ID TRY1_DROME STANDARD; PRT; 252 AA.
AC P52905; Q9V5Y6;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin lota precursor (EC 3.4.21.4).
GN IOFA-TRY OR TRY-IOFA OR CG7754.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RA Gao L., Wang S., Hickey D.A.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Amanatides P.G., Scherer S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Adams M.D., Celnikier S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe C., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bencos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Stroh R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U04853; AAF58655.1;
 DR EMBL; AE003826; AAF58655.1;
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.112;
 DR FlyBase; FBgn0015001; Iota-Try.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Zymogen; Signal; Multigene family.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 27 ACTIVATION PEPTIDE.
 FT CHAIN 28 252 TRYPSIN_IOTA.
 FT ACT_SITE 68 68 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 113 113 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 206 206 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 53 69 BY SIMILARITY.
 FT DISULFID 175 193 BY SIMILARITY.
 FT DISULFID 202 226 BY SIMILARITY.
 FT SITE 200 200 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 252 AA; 26612 MW; 044DEDA4700910C2 CRC64;
 Query Match 94.0%; Score 47; DB 1; Length 252;
 Best Local Similarity 80.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CXGDSGGPXV 10
 | | | | | | |
 DB 202 CTGDSGGPLV 211
 RESULT 2
 CTRL_HALRU STANDARD; PRT; 254 AA.
 ID CTRL_HALRU STANDARD; PRT; 254 AA.
 AC P35003;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chymotrypsin-like serine proteinase precursor (EC 3.4.21.-).
 OS *Halictus rufescens* (California red abalone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
 OC Halitidae; Halictis.
 OX NCBI_TaxID=6454;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE=Intestine;
 RX MEDLINE=93343624; PubMed=8342947;
 RA Grope J.C., Morse D.E.;
 RT "Molluscan chymotrypsin-like protease: structure, localization, and
 RT substrate specificity.";
 RL Arch. Biochem. Biophys. 305:159-169(1993).
 CC -1- FUNCTION: SPECIFICITY SIMILAR TO CHYMOTRYPSIN.

CC -1- ENZYME REGULATION: ACTIVATED BY AN AUTOCATALYTIC MECHANISM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE DISTAL QUARTER
 CC OF THE INTESTINE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X71438; CAA50572.1;
 DR EMBL; S35585; S35585.
 DR PIR; S32750; S32750.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.121;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Glycoprotein; Signal; Zymogen.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 23 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 24 234 CHYMOTRYPSIN-LIKE SERINE PROTEINASE.
 FT ACT_SITE 117 117 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 53 69 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 146 218 BY SIMILARITY.
 FT DISULFID 181 199 BY SIMILARITY.
 FT DISULFID 208 233 BY SIMILARITY.
 SQ SEQUENCE 254 AA; 27250 MW; ADA9A8A22BEFCEC CRC64;
 Query Match 94.0%; Score 47; DB 1; Length 254;
 Best Local Similarity 80.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CXGDSGGPXV 10
 | | | | | | |
 DB 208 CSGDSGGPLV 217
 RESULT 3
 KLR3_HUMAN STANDARD; PRT; 261 AA.
 ID KLR3_HUMAN STANDARD; PRT; 261 AA.
 AC P07288; O16272;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Prostate specific antigen precursor (EC 3.4.21.77) (PSA) (Gamma-
 DE seminoprotein) (kallikrein 3) (Semenogelase) (Seminin) (P-30 antigen).
 GN KLR3 OR APS.
 OS *Homo sapiens* (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=89183632; PubMed=2467258;
 RA Digby M.R., Zhang X.Y., Richard R.I.;
 RT "Human prostate specific antigen (PSA) gene: structure and linkage to
 RT the kallikrein-like gene, hCK-1.";
 RL Nucleic Acids Res. 17:2137-2137(1989).
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE-89282407; PubMed-2471958;
 RA Klobbeck H.-G., Combratio G., Schulz P., Arbusow V., Fittler F.;
 RT "Genomic sequence of human prostate specific antigen (PSA).";
 RL Nucleic Acids Res. 17:3981-3981(1989).
 [3]
 RN SEQUENCE FROM N.A.
 RP TISSUE-Leukocyte;
 RC MEDLINE-89302090; PubMed-2472789;
 RX Lundwall A.;
 RA "Characterization of the gene for prostate-specific antigen, a human
 glandular kallikrein.";
 RT Biochem. Biophys. Res. Commun. 161:1151-1159(1989).
 [4]
 RN SEQUENCE FROM N.A.
 RP TISSUE-Prostate;
 RC MEDLINE-87190978; PubMed-2436946;
 RX Lundwall A., Lilja H.;
 RA "Molecular cloning of human prostate specific antigen cDNA.";
 RT FEBS Lett. 214:317-322(1987).
 [5]
 RN SEQUENCE FROM N.A.
 RP TISSUE-Prostate;
 RC MEDLINE-89246551; PubMed-2470373;
 RX Henttu P., Viikio P.;
 RA "cDNA coding for the entire human prostate specific antigen shows
 high homologies to the human tissue kallikrein genes.";
 RT Biochem. Biophys. Res. Commun. 160:903-910(1989).
 [6]
 RN SEQUENCE FROM N.A.
 RP TISSUE-Prostate;
 RC MEDLINE-89165891; PubMed-2466464;
 RX Triegman P.H.J., Vlietstra R.J., van der Korput J.A.G.M., Romijn J.C.,
 RA "Characterization of the prostate-specific antigen gene: a novel
 human kallikrein-like gene.";
 RT Biochem. Biophys. Res. Commun. 159:95-102(1989).
 [7]
 RN SEQUENCE FROM N.A.
 RP TISSUE-Prostate;
 RC Monne M.M., Moreno J.M., Mele C.M., Mulholland G.M., Gomella L.G.;
 RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 [8]
 RN SEQUENCE FROM N.A.
 RP MEDLINE-20510030; PubMed-11054574;
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
 MOSS P., Paepfer B., Wang K.;
 RT "Sequencing and expression analysis of the serine protease gene
 cluster located in chromosome 19q13 region.";
 RL Gene 257:119-130(2000).
 [9]
 RN SEQUENCE FROM N.A.
 RP TISSUE-Prostate;
 RC Strauberg R.;
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 [10]
 RN SEQUENCE OF 1-176 FROM N.A.
 RX MEDLINE-95079406; PubMed-7527295;
 RA Monne M., Croce C.M., Yu H., Diamandis E.P.;
 RT "Molecular characterization of prostate-specific antigen messenger RNA
 expressed in breast tumors.";
 RL Cancer Res. 54:6344-6347(1994).
 [11]
 RN SEQUENCE OF 17-261 FROM N.A.
 RX MEDLINE-86289366; PubMed-2456523;
 RA Schulz P., Stucka R., Feldmann H., Combratio G., Klobbeck H.-G.,
 RA Fittler F.;
 RT "Sequence of a cDNA clone encompassing the complete mature human
 prostate specific antigen (PSA) and an unspliced leader sequence.";
 RL Nucleic Acids Res. 16:6226-6226(1988).
 [12]
 RN SEQUENCE OF 25-261.
 RX MEDLINE-86205857; PubMed-2422647;

RA Watt K.W.K., Lee P.J., M'Timkulu T., Chan W.P., Loo R.;
 RT "Human prostate-specific antigen: structural and functional
 similarity with serine proteases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3166-3170(1986).
 [13]
 RN SEQUENCE OF 25-261.
 RX MEDLINE-86082806; PubMed-3691515;
 RA Schaller J., Akiyama K., Tsuda R., Hara M., Marti T., Rickli E.E.;
 RT "Isolation, characterization and amino-acid sequence of gamma-
 seminoprotein, a glycoprotein from human seminal plasma.";
 RL Eur. J. Biochem. 170:111-120(1987).
 [14]
 RN 3D-STRUCTURE MODELING.
 RP MEDLINE-95218633; PubMed-7535613;
 RX Villoutreix B.O., Getzoff E.D., Griffin J.H.;
 RA "A structural model for the prostate disease marker, human prostate-
 specific antigen.";
 RT Protein Sci. 3:2033-2044(1994).
 [15]
 RN 3D-STRUCTURE MODELING.
 RP MEDLINE-98427950; PubMed-9751643;
 RX Combs G.S., Bergstrom R.C., Pellequer J.L., Baker S.I., Navre M.,
 RA Smith M.M., Tainer J.A., Madison E.L., Corey D.R.;
 RT "Substrate specificity of prostate-specific antigen (PSA).";
 RL Chem. Biol. 5:475-488(1998).
 CC -1- FUNCTION: PRESUMABLY HYDROLYZE THE HIGH MOLECULAR MASS SEMINAL
 CC VESICLE PROTEIN THUS LEADING TO THE LIQUEFACTION OF THE SEMINAL
 CC COAGULUM.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X14810; CAA32915.1; -;
 DR EMBL; X13940; CAA32123.1; -;
 DR EMBL; X13941; CAA32124.1; ALT_SEQ.
 DR EMBL; X13942; CAB46487.1; -;
 DR EMBL; X13943; CAA32126.1; -;
 DR EMBL; X13944; CAA32127.1; -;
 DR EMBL; X05332; CAA28947.1; -;
 DR EMBL; M27730; -; NOT_ANNOTATED_CDS.
 DR EMBL; M27274; AAA60192.1; -;
 DR EMBL; M26663; AAA58802.1; -;
 DR EMBL; U17040; AAA56764.1; -;
 DR EMBL; AF243527; AAG33355.1; -;
 DR EMBL; BC005307; AAG05307.1; -;
 DR EMBL; S75755; AAD14185.1; ALT_INIT.
 DR EMBL; M24543; AAA60193.1; ALT_SEQ.
 DR PIR; S03604; S03604.
 DR PIR; A26757; A26757.
 DR PIR; A32297; A32297.
 DR PIR; A32423; A32423.
 DR PDB; IPFA; 26-JAN-95.
 DR PDB; 2PSA; 16-MAR-99.
 DR MEROPS; S01.162; -;
 DR Genew; HGNC:6364; KLK3.
 DR MIM; 176820; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Glycoprotein; Antigen; Zymogen; Signal;
 KW 3D-structure.

FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 24 ACTIVATION PEPTIDE.
 FT CHAIN 25 261 PROSTATE SPECIFIC ANTIGEN.
 FT ACT_SITE 65 65 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 69 69 N-LINKED (GLCNAC...).
 FT DISULFID 31 173 BY SIMILARITY.
 FT DISULFID 50 66 BY SIMILARITY.
 FT DISULFID 152 219 BY SIMILARITY.
 FT DISULFID 184 198 BY SIMILARITY.
 FT DISULFID 209 234 BY SIMILARITY.
 FT CONFLICT 64 64 A -> T (IN REF. 11).
 FT CONFLICT 73 73 I -> T (IN REF. 5).
 FT CONFLICT 86 86 T -> I (IN REF. 5).
 FT CONFLICT 94 94 H -> T (IN REF. 12).
 FT CONFLICT 136 136 V -> M (IN REF. 11).
 FT CONFLICT 165 168 FLTP -> HLYDQM (IN REF. 12).
 FT CONFLICT 175 175 D -> Q (IN REF. 12).
 FT CONFLICT 175 175 D -> P (IN REF. 5).
 FT CONFLICT 184 184 C -> Q (IN REF. 5).
 FT CONFLICT 260 260 N -> D (IN REF. 5).
 SQ SEQUENCE 261 AA; 28741 MW; AE9E732AF872141A CRC64;

Query Match 94.0%; Score 47; DB 1; Length 261;
 Best Local Similarity 80.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 I I I I I I I I
 DB 209 CSGDSGGPLV 218

RESULT 4
 ID KIK3_MACMU STANDARD; PRT; 261 AA.
 AC P33619;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Prostate specific antigen precursor (EC 3.4.21.35) (PSA) (Gamma-
 seminoprotein) (Kallikrein 3).
 GN KIK3 OR APS.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93363642; PubMed=7689340;
 RA Gauthier E.R., Chapdelaine P., Tremblay R.R., Dube J.Y.;
 RT "Characterization of rhesus monkey prostate specific antigen cDNA.";
 RL Biochem. Biophys. Acta 1174:207-210(1993).
 CC -1- FUNCTION: PSA PRESUMABLY HYDROLYZE THE HIGH MOLECULAR MASS
 SEMINAL VESICLE PROTEIN THUS LEADING TO THE LIQUEFACTION OF
 THE SEMINAL COAGULUM. PSA IS A GLANDULAR KALLIKREIN.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X73560; CAA51957.1;
 CC PIR: S34239; S34239.
 CC PIR: S35711; S35711.
 CC HSP: P07288; IPFA.
 CC MEROPS: S01.162; -.

DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser-protease_Try.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYPSIN_SPC; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Glycoprotein; Antigen; Zymogen; Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 24 ACTIVATION PEPTIDE.
 FT CHAIN 25 261 PROSTATE SPECIFIC ANTIGEN.
 FT ACT_SITE 65 65 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 31 173 BY SIMILARITY.
 FT DISULFID 50 66 BY SIMILARITY.
 FT DISULFID 152 219 BY SIMILARITY.
 FT DISULFID 184 198 BY SIMILARITY.
 FT DISULFID 209 234 BY SIMILARITY.
 FT CARBOHYD 102 102 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 261 AA; 28816 MW; 8525B14B15967ESC CRC64;

Query Match 94.0%; Score 47; DB 1; Length 261;
 Best Local Similarity 80.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 I I I I I I I I
 DB 209 CSGDSGGPLV 218

RESULT 5
 ID CRAR_HUMAN STANDARD; PRT; 699 AA.
 AC P48740; O95570; Q9UF09;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Complement-activating component of Ra-reactive factor precursor
 DE (EC 3.4.21.-) (Ra-reactive factor serine protease p100) (RaRF)
 DE (Mannan-binding lectin serine protease 1) (Mannose-binding protein
 associated serine protease) (MASP-1).
 GN MASP1 OR CRARF OR CRARF1 OR PRSS5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94059062; PubMed=8240317;
 RA Takada F., Takayama Y., Hatause H., Kawakami M.;
 RT "A new member of the C1s family of complement proteins found in a
 RT bactericidal factor, Ra-reactive factor, in human serum.";
 RL Biochem. Biophys. Res. Commun. 196:1003-1009(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal liver;
 RX MEDLINE=94289349; PubMed=8018603;
 RA Sato T., Endo Y., Matsushita M., Fujita T.;
 RT "Molecular characterization of a novel serine protease involved in
 RT activation of the complement system by mannose-binding protein.";
 RL Int. Immunol. 6:665-669(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=97079701; PubMed=8921412;
 RA Endo Y., Sato T., Matsushita M., Fujita T.;
 RT "Exon structure of the gene encoding the human mannose-binding
 RT protein-associated serine protease light chain: comparison with
 RT complement C1r and C1s genes.";
 RL Int. Immunol. 8:1355-1358(1996).

```

RN SEQUENCE FROM N.A.
RX MEDLINE-99402590; PubMed-10475605;
RA Takayama Y., Takada F., Nowatari M., Kawakami M., Matsui-ura N.;
RT "Gene structure of the p100 serine-protease component of the human Ra-
RL reactive factor";
RL Mol. Immunol. 36:505-514(1999).
CC -|- FUNCTION: COMPONENT OF THE BACTERICIDAL RA-REACTIVE FACTOR RARF
CC WHICH SPECIFICALLY BINDS TO RA AND R2 POLYSACCHARIDES EXPRESSED BY
CC CERTAIN ENTEROBACTERIA. IT TRIGGERS THE ACTIVATION OF COMPLEMENT
CC CASCADE BY ACTIVATING THE C4 AND C2 COMPONENTS. IT ACTIVATES THE
CC C4 COMPONENT BY CLEAVING THE ALPHA-CHAIN OF C4.
CC -|- SUBUNIT: RARF CONSISTS OF A COMPLEMENT-ACTIVATING COMPONENT
CC (CRARF) AND A POLYSACCHARIDE-BINDING (MANNOSE-BINDING) COMPONENT.
CC CRARF IS AN HETERODIMER OF A HEAVY (P70) AND A LIGHT CHAIN (P29)
CC LINKED BY A DISULFIDE BOND.
CC -|- DOMAIN: CRARF HAS A MODULE ORGANIZATION SIMILAR TO CIR AND C1S.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -|- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -|- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17525; BAA04477.1; -
DR EMBL; D28593; BAA05928.1; -
DR EMBL; D61695; BAA34864.1; -
DR EMBL; AB010822; BAA34864.1; JOINED.
DR EMBL; AB010813; BAA34864.1; JOINED.
DR EMBL; AB010814; BAA34864.1; JOINED.
DR EMBL; AB010815; BAA34864.1; JOINED.
DR EMBL; AB010816; BAA34864.1; JOINED.
DR EMBL; AB010817; BAA34864.1; JOINED.
DR EMBL; AB010818; BAA34864.1; JOINED.
DR EMBL; AB010819; BAA34864.1; JOINED.
DR EMBL; AB010820; BAA34864.1; JOINED.
DR EMBL; AB010821; BAA34864.1; JOINED.
DR EMBL; D61690; BAA34864.1; JOINED.
DR EMBL; D61691; BAA34864.1; JOINED.
DR EMBL; D61692; BAA34864.1; JOINED.
DR EMBL; D61693; BAA34864.1; JOINED.
DR EMBL; D61694; BAA34864.1; JOINED.
DR EMBL; AB007617; BAA89206.1; -
DR EMBL; AB007602; BAA89206.1; JOINED.
DR EMBL; AB007603; BAA89206.1; JOINED.
DR EMBL; AB007604; BAA89206.1; JOINED.
DR EMBL; AB007605; BAA89206.1; JOINED.
DR EMBL; AB007606; BAA89206.1; JOINED.
DR EMBL; AB007607; BAA89206.1; JOINED.
DR EMBL; AB007608; BAA89206.1; JOINED.
DR EMBL; AB007609; BAA89206.1; JOINED.
DR EMBL; AB007610; BAA89206.1; JOINED.
DR EMBL; AB007611; BAA89206.1; JOINED.
DR EMBL; AB007612; BAA89206.1; JOINED.
DR EMBL; AB007613; BAA89206.1; JOINED.
DR EMBL; AB007614; BAA89206.1; JOINED.
DR EMBL; AB007615; BAA89206.1; JOINED.
DR EMBL; AB007616; BAA89206.1; JOINED.
DR HSSP; P00736; IAPQ.
DR MEROPS; S01.198; -
DR Genew; HGNC:6901; MASPL.
DR MIM; 600521; -
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.

```

```

DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00084; sush1; 2.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00431; CUB; 2.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00032; CCP; 2.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Complement pathway; Serine protease; Protease;
KW Glycoprotein; Sush1; Repeat; Signal; EGF-like domain; Hydroxylation.
FT SIGNAL 1 19
FT CHAIN 20 699
FT CHAIN 20 448
FT CHAIN 449 699
FT DOMAIN 20 138
FT DOMAIN 139 182
FT DOMAIN 185 297
FT DOMAIN 300 363
FT DOMAIN 366 433
FT DOMAIN 449 699
FT ACT_SITE 490 490
FT ACT_SITE 552 552
FT ACT_SITE 646 646
FT MOD_RES 159 159
FT DISULFID 73 91
FT DISULFID 143 157
FT DISULFID 153 166
FT DISULFID 168 181
FT DISULFID 185 212
FT DISULFID 242 260
FT DISULFID 301 349
FT DISULFID 329 362
FT DISULFID 367 414
FT DISULFID 397 432
FT DISULFID 436 572
FT DISULFID 614 631
FT DISULFID 642 672
FT CARBOHYD 49 49
FT CARBOHYD 178 178
FT CARBOHYD 385 385
FT CARBOHYD 407 407
FT CONFLICT 235 235
FT CONFLICT 285 285
FT CONFLICT 499 499
FT CONFLICT 499 499
FT CONFLICT 527 527
FT CONFLICT 543 543
FT CONFLICT 552 552
FT CONFLICT 643 643
SQ SEQUENCE 699 AA; 79258 MW; ADD9697AE6AB01B5 CRC64;
Query Match 94.0%; Score 47; DB 1; Length 699;
Best Local Similarity 80.0%; Pred. No. 0.37;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXGDSGGPVX 10
DB 642 CAGDSGGPMV 651
RESULT 6
CRAR_MOUSE

```


ID CRAR_MOUSE STANDARD; PRT; 704 AA.
 AC P98064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Complement-activating component of Ra-reactive factor precursor
 DE (EC 3.4.21.-) (Ra-reactive factor serine protease p100) (RaRF)
 DE (Mannan-binding lectin serine protease 1).
 DE MASPI OR CRARF.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-BALB/C; TISSUE=Liver;
 RX MEDLINE=94179811; PubMed=8133044;
 RA Takayama Y., Takada F., Takahashi A., Kawakami M.;
 RT "A 100-kDa protein in the C4-activating component of Ra-reactive
 RT factor is a new serine protease having module organization similar to
 RT C1r and C1s.";
 RL J. Immunol. 152:2308-2316(1994).
 RN [2]
 RP SEQUENCE OF 465-704 FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-BALB/C; TISSUE=Liver;
 RX MEDLINE=93176166; PubMed=8439319;
 RA Takahashi A., Takayama Y., Hattuse H., Kawakami M.;
 RT "Presence of a serine protease in the complement-activating component
 RT of the complement-dependent bactericidal factor, RaRF, in mouse
 RT serum.";
 RL Biochem. Biophys. Res. Commun. 190:681-687(1993).
 CC -1- FUNCTION: COMPONENT OF THE BACTERICIDAL RA-REACTIVE FACTOR RARF
 CC WHICH SPECIFICALLY BINDS TO RA AND R2 POLYSACCHARIDES EXPRESSED BY
 CC CERTAIN ENTEROBACTERIA. IT TRIGGERS THE ACTIVATION OF COMPLEMENT
 CC CASCADE BY ACTIVATING THE C4 AND C2 COMPONENTS. IT ACTIVATES THE
 CC C4 COMPONENT BY CLEAVING THE ALPHA-CHAIN OF C4.
 CC -1- SUBUNIT: RARF CONSISTS OF A COMPLEMENT-ACTIVATING COMPONENT
 CC (CRARF) AND A POLYSACCHARIDE-BINDING (MANNOSE-BINDING) COMPONENT.
 CC CRARF IS AN HETERODIMER OF A HEAVY (P70) AND A LIGHT CHAIN (29)
 CC LINKED BY A DISULFIDE BOND.
 CC -1- TISSUE SPECIFICITY: LIVER.
 CC -1- DOMAIN: CRARF HAS A MODULE ORGANIZATION SIMILAR TO C1R AND C1S.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D16492; BAA03944.1;
 CC HSP: P00736; IAPQ.
 CC MEROPS: S01.198; -.
 CC MGI: 88492; Maspl.
 CC InterPro: IPR000152; Asx_hydroxyl.
 CC InterPro: IPR000859; CUB_domain.
 CC InterPro: IPR001314; Chymotrypsin.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR001881; EGF_CA.
 CC InterPro: IPR001254; Ser-protease_Try.
 CC InterPro: IPR000436; Sush1_SCR_CCP.
 CC Pfam: PF00084; sush1; 2.
 CC Pfam: PF00089; trypsin; 1.
 CC Pfam: PF00431; CUB; 2.
 CC PRINTS: PR00722; CHYMOTRYPSIN.
 CC SMART: SM00032; CCP; 2.
 CC SMART: SM00042; CUB; 2.
 CC SMART: SM00179; EGF_CA; 1.

DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS01186; EGF_CA; 1.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Complement pathway; Serine protease; Protease;
 KW Glycoprotein; Sush1; Repeat; Signal; EGF-like domain; Hydroxylation.
 FT SIGNAL 1 24
 FT CHAIN 25 704 COMPLEMENT-ACTIVATING COMPONENT OF
 FT RA-REACTIVE FACTOR.
 FT CHAIN 25 453 70 KDA CHAIN OF P100 (P70).
 FT CHAIN 454 704 29 KDA CHAIN OF P100 (P29).
 FT DOMAIN 25 143 CUB 1.
 FT DOMAIN 144 187 EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 190 302 CUB 2.
 FT DOMAIN 305 368 SUSHI 1.
 FT DOMAIN 371 438 SUSHI 2.
 FT DOMAIN 454 704 SERINE PROTEASE.
 FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 557 557 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 651 651 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT MOD_RES 164 164 HYDROXYLATION (POTENTIAL).
 FT DISULFID 78 96 POTENTIAL.
 FT DISULFID 148 162 POTENTIAL.
 FT DISULFID 158 171 POTENTIAL.
 FT DISULFID 173 186 POTENTIAL.
 FT DISULFID 190 217 POTENTIAL.
 FT DISULFID 247 265 POTENTIAL.
 FT DISULFID 306 354 POTENTIAL.
 FT DISULFID 334 367 POTENTIAL.
 FT DISULFID 372 419 POTENTIAL.
 FT DISULFID 402 437 POTENTIAL.
 FT DISULFID 441 577 INTERCHAIN (POTENTIAL).
 FT DISULFID 619 636 POTENTIAL.
 FT DISULFID 647 677 POTENTIAL.
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 704 AA; 79895 MW; 71F44F3012DC67F CRC64;
 Query Match 94.0%; Score 47; DB 1; Length 704;
 Best Local Similarity 80.0%; Pred. No. 0.38;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CXGDSGGPKV 10
 Db 547 CAGDSGGPMV 656
 RESULT 7
 LFC_CARRO STANDARD; PRT; 1019 AA.
 ID LFC_CARRO
 AC Q26422;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Limulus clotting factor C precursor (EC 3.4.21.84) (FC).
 OS Carinoscorpius rotundicauda (Southeast Asian horseshoe crab).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
 OC Limulidae; Carinoscorpius.
 OX NCBI_TaxID=6848;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=95268506; PubMed=7538401;
 RA Ding J.L., Navas M.A. III, Ho B.;
 RT "Molecular cloning and sequence analysis of factor C cDNA from the
 RT Singapore horseshoe crab, Carinoscorpius rotundicauda.";
 RL Mol. Mar. Biol. Biotechnol. 4:90-103(1995).

DR EMBL: D90271; BAA14315.1; -
 DR EMBL: D90272; BAA14316.1; -
 DR PIR: A38738; A38738.
 DR PIR: B38738; B38738.
 DR HSP: P00763; IDPO.
 DR MEROPS: S01.219; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR004043; LCCL.dom.
 DR InterPro: IPR001304; LECTIN_C.
 DR InterPro: IPR001254; Ser.protease_Try.
 DR Pfam: PF00059; lectin_C; 1.
 DR Pfam: PF00084; sushi; 5.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: P00722; CHYMOTRYPSIN.
 DR SMART: SM00032; CCP; 5.
 DR SMART: SM00034; CLECT; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS00615; C-TYPE-LECTIN_1; FALSE_NEG.
 DR PROSITE: PS00641; C-TYPE-LECTIN_2; 1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; FALSE_NEG.
 DR PROSITE: PS00820; LCCL; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Hydrolase; Serine protease; Signal; Alternative splicing; Lectin;
 KW Hemolymph clotting; Glycoprotein; Cell adhesion; EGF-like domain;
 KW Sushi; Repeat.
 FT SIGNAL 1 25
 FT CHAIN 26 1019 LIMULUS CLOTTING FACTOR C.
 FT CHAIN 26 690 LIMULUS CLOTTING FACTOR C, HEAVY CHAIN.
 FT CHAIN 691 1019 LIMULUS CLOTTING FACTOR C, LIGHT CHAIN.
 FT CHAIN 691 762 LIMULUS CLOTTING FACTOR C, A CHAIN.
 FT CHAIN 763 1019 LIMULUS CLOTTING FACTOR C, B CHAIN.
 FT DOMAIN 102 137 EGF-LIKE.
 FT DOMAIN 142 195 SUSHI 1.
 FT DOMAIN 200 254 SUSHI 2.
 FT DOMAIN 260 321 SUSHI 3.
 FT DOMAIN 325 421 LCCL.
 FT DOMAIN 436 568 C-TYPE LECTIN.
 FT DOMAIN 576 634 SUSHI 4.
 FT DOMAIN 685 748 SUSHI 5.
 FT DOMAIN 763 1019 SERINE PROTEASE.
 FT ACT_SITE 809 809 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 865 865 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 966 966 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT BINDING 960 960 SUBSTRATE (BY SIMILARITY).
 FT DOMAIN 643 689 PRO-RICH.
 FT DISULFID 106 118 BY SIMILARITY.
 FT DISULFID 112 125 BY SIMILARITY.
 FT DISULFID 127 136 BY SIMILARITY.
 FT DISULFID 436 447 BY SIMILARITY.
 FT DISULFID 464 564 BY SIMILARITY.
 FT DISULFID 538 556 BY SIMILARITY.
 FT DISULFID 794 810 BY SIMILARITY.
 FT DISULFID 932 951 BY SIMILARITY.
 FT DISULFID 962 996 BY SIMILARITY.
 FT CARBOHYD 523 523 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 740 740 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 767 767 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 912 912 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 492 498 LTTTWIG -> TDNVTAT (IN SHORT ISOFORM).
 FT VARSPLIC 499 1019 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 1019 AA; 112346 MW; 58C2864C5715289B CRC64;
 Query Match 94.0%; Score 47; DB 1; Length 1019;
 Best Local Similarity 80.0%; Pred. No. 0.54;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
 | |||||
 Db 962 CSGDSGGPLV 971
 |||||
 RESULT 9
 EL2B_HORSE STANDARD; PRT; 73 AA.
 ID EL2B_HORSE
 AC P37358;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Neutrophil elastase 2B (EC 3.4.21.-) (Proteinase 2B) (Fragments).
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 ON NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Neutrophils;
 RX MEDLINE=94271153; PubMed=7516152;
 RA Dubin A., Potempa J., Travis J.;
 RT "Structural and functional characterization of elastases from horse
 neutrophils".
 RL Biochem. J. 300:401-406(1994).
 CC -!- FUNCTION: MAY BE INVOLVED IN THE DEGRADATION OF CONNECTIVE
 TISSUE IN CHRONIC LUNG DISEASE.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ELASTASE SUBFAMILY.
 DR PIR: S44462; S44462.
 DR HSP: P00766; IGCD.
 DR MEROPS: S01.131; -
 DR InterPro: IPR001254; Ser.protease_Try.
 DR PROSITE: PS0240; TRYPSIN_DOM; PARTIAL.
 DR PROSITE: PS00134; TRYPSIN_HIS; PARTIAL.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 FT NON_CONS 31 32
 FT NON_CONS 56 57
 FT ACT_SITE 64 64 CHARGE RELAY SYSTEM.
 SQ SEQUENCE 73 AA; 7615 MW; AFFB0B330DB69041 CRC64;
 Query Match 92.0%; Score 46; DB 1; Length 73;
 Best Local Similarity 80.0%; Pred. No. 0.06;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CXGDSGGPXV 10
 | |||||
 Db 60 CFGDSGGPLV 69
 |||||
 RESULT 10
 PRTC_CANFA STANDARD; PRT; 157 AA.
 ID PRTC_CANFA
 AC Q28278;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)
 DE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).
 GN PROC.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 ON NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94318474; PubMed=8043441;
 RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
 RT "A comparative study of partial primary structures of the catalytic
 region of mammalian protein C".
 RL Br. J. Haematol. 86:590-600(1994).
 CC -!- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT

REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
-1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va and VIIIA.
-1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D43751; BAA07808.1; -
DR HSSP; P04070; LPCU.
DR MEROPS; S01.218; -
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
FT NON_TER 1 1
FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
FT ACT_SITE 125 125 CHARGE RELAY SYSTEM.
FT DISULFID 96 110 BY SIMILARITY.
FT DISULFID 121 149 BY SIMILARITY.
FT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17262 MW; E8B1BACF49220DFB CRC64;

Query Match 92.0%; Score 46; DB 1; Length 157;
Best Local Similarity 80.0%; Pred. No. 0.13; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 2;

QY 1 CXGDSGGPXY 10
DB 121 CEGDSGGPMV 130

RESULT 11
PRTC_CAPHI ID PRTC_CAPHI STANDARD; PRT; 157 AA.
AC Q28315;
DT 15-DEC-1998 (Rel. 37; Created)
DT 15-DEC-1998 (Rel. 37; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoproteolysin IIA)
DE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).
GN PROC.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94318474; PubMed=8043441;
RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RT "A comparative study of partial primary structures of the catalytic region of mammalian protein C.";
RL Br. J. Haematol. 86:590-600(1994).
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va and VIIIA.
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D43752; BAA07809.1; -
DR HSSP; P04070; LPCU.
DR MEROPS; S01.218; -
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
FT NON_TER 1 1
FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
FT ACT_SITE 125 125 CHARGE RELAY SYSTEM.
FT DISULFID 96 110 BY SIMILARITY.
FT DISULFID 121 149 BY SIMILARITY.
FT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17251 MW; B89790F9954B610A CRC64;

Query Match 92.0%; Score 46; DB 1; Length 157;
Best Local Similarity 80.0%; Pred. No. 0.13; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 2;

QY 1 CXGDSGGPXY 10
DB 121 CEGDSGGPMV 130

RESULT 12
PRTC_FELCA ID PRTC_FELCA STANDARD; PRT; 157 AA.
AC Q28412;
DT 15-DEC-1998 (Rel. 37; Created)
DT 15-DEC-1998 (Rel. 37; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoproteolysin IIA)
DE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).
GN PROC.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94318474; PubMed=8043441;
RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RT "A comparative study of partial primary structures of the catalytic region of mammalian protein C.";
RL Br. J. Haematol. 86:590-600(1994).
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va and VIIIA.
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

or send an email to license@lsb-sib.ch).
CC EMBL; D43750; BAA07807.1; -
CC DR HSPSP; P04070; IPCU.
CC DR MEROPS; S01.218; -
CC DR InterPro; IPR001254; Ser_protease_Try.
CC DR SMART; SM00020; Tryp_SPC; 1.
CC DR PROSITE; PS00240; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
CC FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
CC FT ACT_SITE 125 125 CHARGE RELAY SYSTEM.
CC FT DISULFID 96 110 BY SIMILARITY.
CC FT DISULFID 121 149 BY SIMILARITY.
CC FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT NON_TER 157 157
CC SQ SEQUENCE 157 AA; 17142 MW; FDC5BE2ECA74B4 CRC64;

Query Match 92.0%; Score 46; DB 1; Length 157;
Best Local Similarity 80.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
DB 121 CEGDSGGPMV 130

RESULT 13
PRTC_HORSE
ID PRTC_HORSE STANDARD; PRT; 157 AA.
AC Q2B380;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)
DE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).
GN PROC.
OS Equus caballus (Horse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OC NCBI_Taxid=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94318474; PubMed-8043441;
RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RT "A comparative study of partial primary structures of the catalytic
region of mammalian protein C";
RL Br. J. Haematol. 86:590-600(1994).
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC and VIIa.
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL; D43753; BAA07810.1; -
CC DR HSPSP; P04070; IPCU.
CC DR MEROPS; S01.218; -
CC DR InterPro; IPR001254; Ser_protease_Try.
CC PFam; PF00089; trypsin; 1.

```

```

DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
DR FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
DR FT ACT_SITE 125 125 CHARGE RELAY SYSTEM.
DR FT DISULFID 96 110 BY SIMILARITY.
DR FT DISULFID 121 149 BY SIMILARITY.
DR FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
DR FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
DR FT NON_TER 157 157
DR SQ SEQUENCE 157 AA; 17200 MW; BFA6EA045C3580 CRC64;

Query Match 92.0%; Score 46; DB 1; Length 157;
Best Local Similarity 80.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
DB 121 CEGDSGGPMV 130

RESULT 14
PRTC_MACMU
ID PRTC_MACMU STANDARD; PRT; 161 AA.
AC Q2B506;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)
DE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).
GN PROC.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OC NCBI_Taxid=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94318474; PubMed-8043441;
RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RT "A comparative study of partial primary structures of the catalytic
region of mammalian protein C";
RL Br. J. Haematol. 86:590-600(1994).
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC and VIIa.
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL; D43754; BAA07811.1; -
CC DR HSPSP; P04070; IPCU.
CC DR MEROPS; S01.218; -
CC DR InterPro; IPR001254; Ser_protease_Try.
CC PFam; PF00089; trypsin; 1.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
CC FT NON_TER 157 157

```

```

FT ACT_SITE 26 26 CHARGE RELAY SYSTEM.
FT ACT_SITE 129 129 CHARGE RELAY SYSTEM.
FT DISULFID 100 114 BY SIMILARITY.
FT DISULFID 125 153 BY SIMILARITY.
FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 161 161
SQ SEQUENCE 161 AA; 17770 MW; 27D78F185B2FCC69 CRC64;

Query Match 92.0%; Score 46; DB 1; Length 161;
Best Local Similarity 80.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
Db 125 CEGDSGGPMV 134

RESULT 15
TRY3_LUCCU STANDARD; PRT; 165 AA.
AC P35043;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trypsin alpha-3 (EC 3.4.21.4) (Fragment)
OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Oestroidea; Calliphoridae; Lucilia.
OX NCBI_TaxID=7375;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95202118; PubMed=7894748;
RA Casu R.E., Jarney J.M., Elvin C.M., Elsemann C.H.;
RT "Isolation of a trypsin-like serine protease gene family from the
RT sheep blowfly Lucilia cuprina.";
RL Insect Mol. Biol. 3:159-170(1994).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L15632; AAA65931.1;
CC HSP; P00763; IDPO.
CC DR MEROPS; S01.112;
CC DR InterPro; IPR001254; Ser_protease_Try.
CC DR Pfam; PF00089; trypsin; 1.
CC DR SMART; SM00020; TRYD_SPC; 1.
CC DR PROSITE; PS00240; TRYPsin_DOM; 1.
CC DR PROSITE; PS00134; TRYPsin_HIS; PARTIAL.
CC DR PROSITE; PS00135; TRYPsin_SER; 1.
CC KW Hydrolase; Serine protease; Multigene family.
CC FT NON_TER 1
CC FT ACT_SITE 26 26 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT ACT_SITE 119 119 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC FT DISULFID 89 106 BY SIMILARITY.
CC FT DISULFID 115 139 BY SIMILARITY.
CC FT SITE 113 113 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
CC SQ SEQUENCE 165 AA; 16569 MW; 26160B1AFF80F1CD CRC64;

Query Match 92.0%; Score 46; DB 1; Length 165;
Best Local Similarity 80.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10

```

Db 115 CQDSGGPLV 124

Search completed: May 12, 2003, 15:36:38
Job time : 6.8 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:32:37 ; Search time 22.6 Seconds
(without alignments)
91.171 Million cell updates/sec

Title: US-09-909-348-2

Perfect score: 50

Sequence: 1 CXGDSGGPXV 10

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:**

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organelle:**
- 9: sp_phase:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**
- 15: sp_virus:**
- 16: sp_bacteriap:**
- 17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	96.0	258	5	Q97399
2	47	94.0	152	16	Q98dl1
3	47	94.0	214	5	Q9VGC0
4	47	94.0	247	5	O18458
5	47	94.0	248	5	O16126
6	47	94.0	248	5	Q9VQ98
7	47	94.0	254	5	Q95UR4
8	47	94.0	259	5	Q9VRS8
9	47	94.0	259	5	Q9VRD0
10	47	94.0	260	5	Q9VRD1
11	47	94.0	262	5	Q9VRS9
12	47	94.0	265	5	O16101
13	47	94.0	266	5	Q9VMA7
14	47	94.0	268	5	O17030
15	47	94.0	271	5	Q9V4W5
16	47	94.0	272	5	Q8SZG4

17	47	94.0	276	5	Q97398
18	47	94.0	276	5	Q9VMX8
19	47	94.0	282	11	Q9D413
20	47	94.0	287	5	Q9NH10
21	47	94.0	287	5	Q9VTV2
22	47	94.0	296	5	Q8T9U6
23	47	94.0	322	11	Q920S2
24	47	94.0	345	5	Q86088
25	47	94.0	362	5	Q9W453
26	47	94.0	366	11	O70170
27	47	94.0	412	5	Q9VVI4
28	47	94.0	477	5	Q9VZT0
29	47	94.0	483	5	O8T8X4
30	47	94.0	580	5	Q9VS87
31	47	94.0	611	5	Q9VZS8
32	47	94.0	698	13	Q9PU71
33	47	94.0	701	11	Q9JJ59
34	47	94.0	774	5	Q9VK10
35	47	94.0	1019	5	O8T9S1
36	47	94.0	1083	5	O26423
37	47	94.0	1190	5	Q9V7E5
38	46	92.0	50	5	Q9GTK6
39	46	92.0	66	11	Q61752
40	46	92.0	73	6	Q9TV90
41	46	92.0	87	11	Q9C078
42	46	92.0	110	11	Q9RLW7
43	46	92.0	117	13	Q9PUF3
44	46	92.0	119	4	Q9NR68
45	46	92.0	138	11	Q9D725

ALIGNMENTS

RESULT 1

Q97399
ID Q97399 PRELIMINARY; PRT; 258 AA.
AC Q97399;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Trypsin precursor.
OS Phaedon cochleariae (Mustard beetle).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Chrysomeloidea; Chrysomelidae;
OC Chrysomelinae; Phaedon.
OX NCBI_TaxID=80249;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=GUT;
RA Girard C., Jouanin L.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
CC EMBL; Y17905; CAA76929.1; -
DR HSP; P00763; IDPO.
DR MEROPS; S01.113; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00089; trypsin; 1
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SEC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Serine protease; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 30 258 TRYPSIN.
SQ SEQUENCE 258 AA; 28070 MW; BBDFAFECB86866C CRC64;

Query Match 96.0%; Score 48; DB 5; Length 258;
Best Local Similarity 80.0%; Pred. No. 0.29;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 CXGDSGGPXV 10
DB 209 CSGDSGGPAV 218

RESULT 2
Q98DL1 PRELIMINARY: PRT; 152 AA.
AC Q98DL1
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein mlr4654.
GN MLR4654.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WAFF303099;
RX MEDLINE-21082930; PubMed-11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Igesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003004; BAB51260.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 152 AA; 16393 MW; E74A9FF9E5567A1F CRC64;

Query Match 94.0%; Score 47; DB 16; Length 152;
Best Local Similarity 80.0%; Pred. No. 0.27;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
DB 26 CGGDSGGPTV 35

RESULT 3
Q9VGC0 PRELIMINARY: PRT; 214 AA.
AC Q9VGC0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG17404 protein.
GN CG17404.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Arkil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balwle R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

```

```

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.-D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hartis N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacile J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF003695; AAF54763.1; -
DR HSSP; P00766; 1CHG.
DR FlyBase; FBgn0038001; CG17404.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 214 AA; 23006 MW; 06905997C4EA64C9 CRC64;

Query Match 94.0%; Score 47; DB 5; Length 214;
Best Local Similarity 80.0%; Pred. No. 0.38;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
DB 160 CSGDSGGPLV 169

RESULT 4
O18458 PRELIMINARY: PRT; 247 AA.
ID O18458
AC O18458;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serine proteinase precursor.
GN SP-II.
OS Heterodera glycines (Soybean cyst nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
OX NCBI_TaxID=51029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98030247; PubMed-9364965;
RA Lilley C.J., Urwin P.E., Atkinson H.J., McPherson M.J.;
RT "Characterization of cDNAs encoding serine proteinases from the
RT soybean cyst nematode Heterodera glycines."
RL Mol. Biochem. Parasitol. 89:195-207(1997).
DR EMBL; Y13907; CAA74205.1; -

```



```

DR HSP: P00763; 1DPO.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 22 247 SERINE PROTEINASE.
SQ SEQUENCE 247 AA; 25586 MW; 3A2B5B2B3BB77222 CRC64;

Query Match 94.0%; Score 47; DB 5; Length 247;
Best Local Similarity 80.0%; Pred. No. 0.44;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
DB 201 CSGDSGGPLV 210

RESULT 5
O16126 PRELIMINARY; PRT; 248 AA.
AC O16126;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Trypsinogen 1 precursor.
GN TRYPI.
OS Boltenia villosa.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pygididae; Boltenia.
OX NCBI_TaxID=63515;
RN [1]
RP SEQUENCE FROM N.A.
RA "The Molecular Evolution of the Vertebrate Trypsinogens.";
RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RL
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AF011897; AAB69563.1; -.
DR HSP: P00763; 1DPO.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPsin.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR Hydrolase; Serine protease; Signal.
FT SIGNAL 1 12 POTENTIAL.
FT CHAIN 13 248 TRYPSIN 1.
SQ SEQUENCE 248 AA; 25872 MW; AC60688998413305 CRC64;

Query Match 94.0%; Score 47; DB 5; Length 248;
Best Local Similarity 80.0%; Pred. No. 0.44;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
DB 200 CQDGGGPAV 209

RESULT 6
Q9VQ98 PRELIMINARY; PRT; 248 AA.
AC Q9VQ98;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CG10882 protein.
GN CG17239.
OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt E.G., Nelson C.R., Mikos G.D.G.,
RA Khril J.F., Agbayani A., An H.-J., Andrews-Fiankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Koldra C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AE003583; AAF51279.1; -.
DR HSP: P00763; 1DPO.
DR FlyBase: FBgn0042186; CG17239.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPsin.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 248 AA; 26744 MW; 07B0FD663F005807 CRC64;

Query Match 94.0%; Score 47; DB 5; Length 248;
Best Local Similarity 80.0%; Pred. No. 0.44;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
DB 194 CSGDSGGPLV 203

RESULT 7
Q95UP4

```

ID Q95UP4 PRELIMINARY; PRT; 254 AA.
AC Q95UP4; 01-DEC-2001 (TREMBlrel. 13, Created)
DT 01-DEC-2001 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Serine protease Ssp3.
GN Ssp3.
OS Stomoxys calcitrans (Stable fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Muscoidea; Muscidae; Stomoxys.
OX NCBI_TaxID=35570;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton J.V., Munks R.J.L., Lehane S.M., Lehane M.J.;
RT "Association of Intestinal Defensins with a Novel Serine Protease in
RT The Blood Sucking Fly Stomoxys calcitrans."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY044834; AAK98781.1; -;
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 254 AA; 27521 MW; 99AEB8B1481FA439 CRC64;
Query Match 94.0%; Score 47; DB 5; Length 254;
Best Local Similarity 80.0%; Pred. No. 0.45;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXGDSGGPXPV 10
Db | ||||| |
Db 206 CRGDSGGPVP 215
RESULT 8
Q9VRS8 PRELIMINARY; PRT; 259 AA.
AC Q9VRS8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE CG6580 protein.
GN CG6580.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burling K.J., Evans C.A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
RA Foshier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matteli B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Morkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., M.G.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AE003564; AAF50712.1; -;
DR HSP; P00763; IDPO.
DR FlyBase; FBgn0035666; CG6580.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 259 AA; 28373 MW; 13249A284038F4F9 CRC64;
Query Match 94.0%; Score 47; DB 5; Length 259;
Best Local Similarity 80.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CXGDSGGPXPV 10
Db | ||||| |
Db 205 CSGDSGGPVP 214
RESULT 9
Q9VRD0 PRELIMINARY; PRT; 259 AA.
AC Q9VRD0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE SER6 protein (RE66795p).
GN SER6 OR CG2071.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burling K.J., Evans C.A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
RA Foshier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Shue B.C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY
DR EMBL; AE003569; AAF50872.1; -;
DR EMBL; AF089642; AAL0380.1; -;
DR HSSP; P00761; 1EPT.
DR FlyBase; FBgn0011834; Ser6.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR SMART; SM00020; TRYp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 259 AA; 28078 MW; 49D5EB6E37A9D4DF CRC64;

Query Match 94.0%; Score 47; DB 5; Length 259;
Best Local Similarity 80.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGPKV 10
Db 210 CNGDSGPAV 219

RESULT 10
Q9VRD1 PRELIMINARY; PRT; 260 AA.
AC Q9VRD1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG1304 protein (REL14171p).
GN CG1304.

OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokov D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY
DR EMBL; AE003569; AAF50871.1; -;
DR EMBL; AY095196; AAM12289.1; -;
DR HSSP; P00761; 1EPT.
DR FlyBase; FBgn0031141; CG1304.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR SMART; SM00020; TRYp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 260 AA; 27860 MW; 0933B34BF20A7675 CRC64;

Query Match 94.0%; Score 47; DB 5; Length 260;
Best Local Similarity 80.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
I | | | | |
Db 210 CNGDSGGPVP 219

RESULT 11

Q9VRS9 PRELIMINARY; PRT; 262 AA.
AC Q9VRS9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE CG10475 protein.
GN CG10475.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Broksstein P., Brattier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,
RA Reihert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhou X., Zhou Q., Zhou L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AE003564; AAF50711.1; -;
DR HSP; P00734; I87X.
DR FlyBase; FBgn0035667; CG10475.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.
RW Hydrolase; Serine protease.
SQ SEQUENCE 262 AA; 28511 MW; 7F0F3F66044082CA CRC64;

Query Match 94.0%; Score 47; DB 5; Length 262;
Best Local Similarity 80.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
I | | | | |
Db 208 CNGDSGGPVP 217

RESULT 12

O16101 PRELIMINARY; PRT; 265 AA.
AC O16101;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Serine protease SER4 precursor.
GN SER4 OR CG8867.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OREGON R;
RX MEDLINE=98146263; PubMed=9473638;
RA Ahrens J.E., Mahoney P.A.;
RT "Isolation and cloning of Ser4, a gene encoding a trypsin-like serine
RT protease in Drosophila melanogaster.";
RL Biochim. Biophys. Acta 1395:141-144(1998).
RE EMBL; AF006639; AAC14351.1; -;
DR HSP; P20231; IAAO
DR FlyBase; FBgn0020906; Ser4.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 37 265 ACTIVE SERINE PROTEASE SER4.
SQ SEQUENCE 265 AA; 28950 MW; 551A92CDE2E28BC8 CRC64;
Query Match 94.0%; Score 47; DB 5; Length 265;
Best Local Similarity 80.0%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CXGDSGGPXV 10
I | | | | |
Db 210 CNGDSGGPVP 219

RESULT 13
Q9VNM7 PRELIMINARY; PRT; 266 AA.
AC Q9VNM7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE SER4 protein.
GN SER4 OR CG8867.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OX Anophelinae
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G3; TISSUE=MIDGUT;
RX MEDLINE=97075119; PubMed=8917545;
RA Dimopoulos G.M., Richman A., della Torre A., Kafatos F.C., Louis C.;
RT "Identification and characterization of differentially expressed cDNAs
of the vector mosquito, Anopheles gambiae";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13066-13071(1996).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
EMBL; Z69978; CAA93818.1; .
DR HSP; P00763; LDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease TRY.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC_1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00334; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 268 AA; 29176 MW; 7BEE8462EF1FD8BE CRC64;

Query Match 94.0%; Score 47; DB 5; Length 268;
Best Local Similarity 80.0%; Pred. NO. 0.47;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps

QY 1 CXGDGGGPXV 10
| | ||||| |
DB 215 CSGDGGGPLV 224

RESULT 15

Q9V4W5 PRELIMINARY; PRT; 271 AA.

AC Q9V4W5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG8579 protein (RE27458p).
GN CG8579.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Ephydroidea; Drosophilidae; Drosophilla.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeell M.D., Zhang Q., Chen L.X.,
RA Brandon K.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Beriman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Butchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garr N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harrijs N.I., Harvey D., Helman T.J., Hernandez J.P., Houck T,
RA Hevner J., Holmberg M., Hornok T.

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzay D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY
 DR EMBL; AE003836; AAF59060.2; -;
 DR EMBL; AY095199; AAMI2292.1; -;
 DR HSP; P00763; IDPO.
 DR FlyBase; FBgn0033319; CG8579.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR01254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 271 AA; 29814 MW; 0A410C925CBE8826 CRC64;

Query Match 94.0%; Score 47; DB 5; Length 271;
 Best Local Similarity 80.0%; Pred. No. 0.48;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CXGSGGPGPV 10
 Db 217 CSGSGGPGPV 226

Search completed: May 12, 2003, 15:38:40
 Job time : 24.6 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:31:22 ; Search time 11.44 Seconds
(without alignments)
46.591 Million cell updates/sec

Title: US-09-909-348-3
Perfect score: 21
Sequence: 1 RGDA 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	4	13	Cell contact inhib
2	21	100.0	4	22	Transport molecule
3	21	100.0	4	23	Human thrombin pep
4	21	100.0	4	23	Thrombin peptide d
5	21	100.0	4	23	Thrombin receptor
6	21	100.0	5	13	Platelet antagonist
7	21	100.0	5	20	Human thrombospond
8	21	100.0	5	22	Thrombin-induced p
9	21	100.0	6	11	Peptide from fibro
10	21	100.0	6	12	Cell attachment pr

11	21	100.0	7	23	ABP48385	zinc finger protei
12	21	100.0	7	23	ABP48394	zinc finger protei
13	21	100.0	7	23	ABP48597	zinc finger protei
14	21	100.0	7	23	ABP48600	zinc finger protei
15	21	100.0	7	23	ABP48603	zinc finger protei
16	21	100.0	7	23	ABP48606	zinc finger protei
17	21	100.0	7	23	ABP48609	zinc finger protei
18	21	100.0	7	23	ABP48672	zinc finger protei
19	21	100.0	7	23	ABP48868	zinc finger protei
20	21	100.0	7	23	ABP48883	zinc finger protei
21	21	100.0	7	23	ABP49111	zinc finger protei
22	21	100.0	7	23	ABP49114	zinc finger protei
23	21	100.0	7	23	ABP49409	zinc finger protei
24	21	100.0	7	23	ABP49436	zinc finger protei
25	21	100.0	7	23	ABP49439	zinc finger protei
26	21	100.0	7	23	ABP49445	zinc finger protei
27	21	100.0	7	23	ABP49448	zinc finger protei
28	21	100.0	7	23	ABP49547	zinc finger protei
29	21	100.0	7	23	ABP49631	zinc finger protei
30	21	100.0	7	23	ABP49634	zinc finger protei
31	21	100.0	7	23	ABP49652	zinc finger protei
32	21	100.0	7	23	ABP49655	zinc finger protei
33	21	100.0	7	23	ABP49679	zinc finger protei
34	21	100.0	7	23	ABP49685	zinc finger protei
35	21	100.0	7	23	ABP50170	zinc finger protei
36	21	100.0	7	23	ABP50194	zinc finger protei
37	21	100.0	7	23	ABP50270	zinc finger protei
38	21	100.0	7	23	ABP50310	zinc finger protei
39	21	100.0	7	23	ABP50315	zinc finger protei
40	21	100.0	7	23	ABP50411	zinc finger protei
41	21	100.0	7	23	ABP50964	zinc finger protei
42	21	100.0	7	23	ABP51010	zinc finger protei
43	21	100.0	8	19	AAW48507	integrin receptor
44	21	100.0	9	12	AAR12332	platelet-aggregati
45	21	100.0	9	12	AAR12333	platelet-aggregati

ALIGNMENTS

RESULT 1
AAR25315
ID AAR25315 standard; peptide; 4 AA.
XX AC AAR25315;
XX DT 17-MAR-1993 (first entry)
XX DE Cell contact inhibitor generic peptide #4.
XX KW Cyclic peptide; cell contact inhibitor; hydrolytic enzyme.
XX OS Synthetic.
XX FH Key
XX FT Modified-site 2
XX FT /label- MeGly
XX PN JP04264097-A.
XX PD 18-SEP-1992.
XX PF 16-FEB-1991; 91JP-0044386.
XX PR 16-FEB-1991; 91JP-0044386.
XX PA (ASAG) ASAHI GLASS CO LTD.
XX DR WPI; 1992-361922/44.
XX PT Peptide derivs. as contact inhibitor for animal cells - comprise
PT synthesised cyclic peptide and have portion of aminoacid sequence
PT of arginine-N-methyl:glycine-aspartic acid

XX Disclosure; Page 3; 6pp; Japanese.

XX The sequences given in AAR25311-19 are cyclic peptides which act as contact inhibitors of animal cells. They are resistant to decomposition by hydrolytic enzymes and can be maintained at high levels of activity for a long period in vivo. The peptides are cyclic and may have 1-16 pref. 1-4 amino acids.

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 13; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
||||

Db 1 RGDA 4

RESULT 2

AA86859

ID AAB86859 standard; peptide; 4 AA.

XX AAB86859;

XX

DT 28-NOV-2001 (first entry)

XX

DE Transport molecule/ligand binding-associated peptide #5.

XX

KW Transport molecule; ligand; cancer treatment; autoimmune disease;

KW Inflammation; infection.

XX Synthetic.

OS

XX WO200168142-A1.

PN

XX

PD 20-SEP-2001.

XX

PF 13-MAR-2001; 2001WO-EP02833.

XX

PR 13-MAR-2000; 2000DE-1012120.

XX

XX (KTB-) KTB TUMORFORSCHUNGS GMBH.

PA

XX Kratz F;

PI

XX WPI; 2001-589998/66.

DR

XX New ligand, comprising therapeutic or diagnostic agent bonded non-covalently with substance having high affinity to transport molecule

PT

PT

XX Disclosure; Page 39; 74pp; German.

XX This invention describes novel ligands which bind to transport molecules, comprising a therapeutic and/or diagnostic agent (A) non-covalently bonded via a linkage cleavable in vivo depending on pH and/or enzymatically with a substance (B) having an association constant KA to a transport molecule of above 10³ M⁻¹, is new. The medicaments are especially useful for the treatment of cancers, autoimmune diseases, acute and chronic inflammation and infections caused by viruses or microorganisms. The diagnostic kits are useful for the detection of these illnesses and for the detection of the transport molecule and/or its distribution in vivo. The ligands have excellent solubility in the medium at the site of action and are easy and inexpensive to convert into adducts, as the interaction with the transport material is physical. AAB86843-AAB86920 represent peptides used to illustrate the method of the invention.

XX

SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 22; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
||||

Db 1 RGDA 4

RESULT 3

AAE20157

ID AAE20157 standard; peptide; 4 AA.

XX

AC AAE20157;

XX

DT 18-JUN-2002 (first entry)

XX

DE Human thrombin peptide.

XX

KW Cartilage growth; cartilage repair; arthritic joint; traumatic injury; non-proteolytically activated thrombin receptor; NPAR; chondrocyte; therapy; implantation; thrombin peptide; human.

KW

XX Homo sapiens.

OS

XX WO200207748-A2.

PN

XX

PD 31-JAN-2002.

XX

PF 19-JUL-2001; 2001WO-US22668.

XX

PR 20-JUL-2000; 2000US-219800P.

XX

XX (TEXA) UNIV TEXAS SYSTEM.

PA

XX Carney DH, Crowther RS, Stiernberg J, Bergmann J;

PI

XX WPI; 2002-268953/31.

DR

XX Stimulating growth and repair of cartilage, useful for treating e.g. arthritis, by local administration of an agonist of non-proteolytically activated thrombin receptor

PT

PT

XX Claim 10; Page 25; 28pp; English.

PS

XX The invention relates to a method of stimulating growth and repair of cartilage. The method involves administering to the site, an agonist of non-proteolytically activated thrombin receptor (NPAR). The method is used in human or veterinary medicine for the treatment of arthritic joints and damage/loss of cartilage caused by traumatic injury. Also chondrocytes may be cultured in presence of NPAR agonist to provide cells for implantation at sites requiring growth/repair of cartilage. The present sequence is human thrombin peptide. The derivatives of thrombin peptide which serves as a NPAR agonist.

CC

XX

SQ Sequence 4 AA;

Query Match 100.0%; Score 21; DB 23; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
||||

Db 1 RGDA 4

RESULT 4

AAU78374

ID AAU78374 standard; Peptide; 4 AA.

XX

AC AAU78374;

XX

DT 18-JUN-2002 (first entry)

XX

DE Thrombin peptide derivative #1.

XX Thrombin; osteopathic; receptor; agonist; bone growth stimulation;
 KW osteoinduction; farm animal; companion animal; laboratory animal;
 KW bone graft; segmental bone gap; bone void; non-union fracture.

XX Synthetic.

XX WO200205836-A2.

XX 24-JAN-2002.

XX 18-JUL-2001; 2001WO-US22641.

XX 19-JUL-2000; 2000US-219300P.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Carney DH, Crowther RS, Simmons DJ, Yang J, Redin WR;

XX WPI; 2002-303796/34.

XX Stimulating bone growth at a site in a subject in need of
 PT osteoinduction, such as a site of bone graft, segmental bone gap, bone
 PT void or non-union structure, by administering agonist of activated
 PT thrombin receptor

XX Claim 9; Page 22; 27pp; English.

XX The invention describes a method of stimulating bone growth at a site
 CC in a subject in need of osteoinduction. The method involves administering
 CC an agonist to stimulate bone growth at a site in a subject (e.g. a farm
 CC animal, companion animal or laboratory animal), in need of
 CC osteoinduction, such as the site in need of a bone graft in a subject, a
 CC segmental bone gap, a bone void or a non-union fracture. This sequence
 CC represents a thrombin peptide derivative obtained from a serine
 CC esterase that can stimulate or activate the non-proteolytically
 CC activated thrombin receptor.

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 23; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4

||||

Db 1 RGDA 4

RESULT 5

AAAM50856
 ID AAAM50856 standard; Peptide: 4 AA.

XX AC AAAM50856;

XX 01-MAY-2002 (first entry)

XX Thrombin receptor binding domain used for cardiac tissue repair.

XX Thrombin receptor binding domain; thrombin; revascularisation;
 KW vascular occlusion; tissue repair; vulnarary; vasotropic; cardiant;
 KW angiogenesis; restenosis; therapy; human.

XX Homo sapiens.

XX WO200204008-A2.

XX 17-JAN-2002.

XX 12-JUL-2001; 2001WO-US21944.

XX 12-JUL-2000; 2000US-217583P.

XX PA

(TEXA) UNIV TEXAS SYSTEM.

XX Carney DH;

XX WPI; 2002-179665/23.

XX Promoting cardiac tissue repair, stimulating revascularisation,
 PT stimulating vascular endothelial cell proliferation, and inhibiting
 PT vascular occlusion by using angiogenic thrombin derivative peptide

XX Claim 2; Page 19; 24pp; English.

XX The present sequence is that of a thrombin receptor binding domain
 CC peptide that is used in a claimed method for promoting cardiac
 CC tissue repair. The method involves administering an angiogenic
 CC thrombin-derived peptide. The peptide comprises the present
 CC thrombin receptor binding domain together with a serine esterase
 CC conserved sequence (see AAM50857), or preferably a peptide (see
 CC AAM50858) which includes both these sequences. The thrombin-derived
 CC peptide is administered during or following cardiac surgery by
 CC injection into cardiac tissue, and may be formulated as a sustained
 CC release formulation. It is used in claimed methods of stimulating
 CC revascularisation, stimulating vascular endothelial cell
 CC proliferation, inhibiting vascular occlusion, and inhibiting
 CC restenosis following balloon angioplasty, in which case the
 CC peptide may be coated onto the catheter.

XX Sequence 4 AA;

Query Match 100.0%; Score 21; DB 23; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4

||||

Db 1 RGDA 4

RESULT 6

AAAR24517
 ID AAAR24517 standard; Protein: 5 AA.

XX AC AAAR24517;

XX 02-DEC-1992 (first entry).

XX Platelet antagonist peptide 4.

XX Clinical effect; antagonist.

XX Synthetic.

XX JP04134096-A.

XX -07-MAY-1992.

XX 21-SEP-1990; 90JP-0253849.

XX 21-SEP-1990; 90JP-0253849.

XX (SEGK) SEIKAGAKU KOGYO CO LTD.

XX WPI; 1992-204525/25.

XX New peptide(s) comprising arginine-glycine-asparagine and
 PT hyaluronic acid - useful as platelet antagonists with higher
 PT activity than arginine-glycine-asparagine-valine

XX Disclosure; Page 5; 10pp; Japanese.

XX The sequences given in AAR24514-8 are peptides which are useful as
 CC platelet antagonists. These peptides have higher activity than the

CC conventional peptide of Arg-Gly-Asp-Val. These peptides have a
 CC clinical effect at a lower dose, dosage is 2.5-5.0 mg/kg/day.

XX Sequence 5 AA;

Query Match 100.0%; Score 21; DB 13; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 DB 2 RGDA 5

RESULT 7

AA17781
 ID AAY17781 standard; peptide; 5 AA.

XX AC

XX AAY17781;

XX 12-AUG-1999 (first entry)

XX Human thrombospondin-1 type III repeat peptide.

XX Human; thrombospondin; HIV; infection; inhibition; chemokine;
 KW contraceptive.

XX Homo sapiens.

OS Synthetic.

XX WO926649-A1.

XX 03-JUN-1999.

XX 24-NOV-1998; 98WO-US24905.

XX 20-MAR-1998; 98US-0078873.

XX 25-NOV-1997; 97US-0066294.

XX (CORR) CORNELL RES FOUND INC.

XX Crombie AR, Laurence JC, Nachman RL;

PI WPI; 1999-370856/31.

XX Suppressing infectivity of human immune deficiency virus

XX Example 2; Page 33; 67pp; English.

CC The present invention describes a method for suppressing infectivity of
 CC human immunodeficiency virus (HIV) by treating the virus, or its target
 CC cell, with a thrombospondin or thrombospondin analogue. Thrombospondin
 CC blocks binding of HIV to its cellular receptors. Thrombospondin or its
 CC analogues can be used to prevent infection by HIV, in both contraceptive
 CC and non-contraceptive compositions/devices. They are already known to
 CC reduce infectivity of some bacteria and protozoa. The present sequence
 CC represents a human thrombospondin-1 type III repeat peptide.

XX Sequence 5 AA;

Query Match 100.0%; Score 21; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 DB 2 RGDA 5

RESULT 8

AA172600

ID AAB72600 standard; Peptide; 5 AA.

XX

AC AAB72600;

XX 09-MAY-2001 (first entry)

XX Thrombin-induced platelet activator antagonist #39.

XX Platelet aggregation inhibitor; thrombin activation inhibitor;

XX protease activated receptor 1; PAR1; platelet activation inhibitor;

XX thrombosis; acute coronary syndrome.

XX Unidentified.

XX WO200112656-A1.

XX 22-FEB-2001.

XX 17-AUG-2000; 2000WO-US40669.

XX 17-AUG-1999; 99US-0375808.

XX (THRO-) THROMGEN INC.

XX Schmaier AH, Hasan AAK;

XX WPI; 2001-226546/23.

XX Inhibiting thrombin activation in human cell expressing protease
 CC activated receptor 1 (PAR1), comprises contacting mixtures of thrombin
 CC and human cell expressing PAR1, with a peptide that inhibits platelet
 CC activation

XX Claim 8; Page 26; 49pp; English.

CC The present invention relates to a method for inhibiting thrombin
 CC activation in a human cell expressing protease activated receptor 1
 CC (PAR1). The method involves using peptides (e.g. the present peptide)
 CC that inhibit platelet activation. The method is useful for preventing
 CC thrombosis and platelet aggregation. The method can be used for patients
 CC with acute coronary syndromes (e.g. crescendo angina, myocardial
 CC infarction) and for individuals who have acute coronary syndromes and
 CC receive percutaneous transluminal coronary angioplasty with an article
 CC stent placement.

XX Sequence 5 AA;

Query Match 100.0%; Score 21; DB 22; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4

DB 1 RGDA 4

RESULT 9

AA172600

ID AAR04871 standard; peptide; 6 AA.

XX AAR04871;

XX 25-SEP-1989 (first entry)

XX Peptide from fibronectin.

XX Fibronectin; cell attachment; cell detachment; fermentation; therapy.
 XX synthetic.
 XX OS
 XX US4879237-A.
 XX 07-NOV-1989.
 XX 24-MAY-1985; 85US-0738078.

XX 24-MAY-1989; 89US-0738078.
 XX (JOLL-) LA JOLLA CANCER RES.
 XX Ruoslahti EI, Hayman EG, Pierschbacher MD;
 XX WPI; 1990-154405/20.
 XX Synthetic peptide(s) from fibronectin- used in control of cell attachment
 XX and detachment
 XX Claim 1; page 10; 13pp; English.
 XX This polypeptide mediates the attachment of animal cells to substrates.
 XX The substrate (I) is contacted with cells and with a soln. contg. this
 XX polypeptide. This attachment can be prevented in addition to detaching
 XX the cells from (I) once attached. Applications are in eg fermentation,
 XX cell line prepn., diagnosis and therapy.
 XX Sequence 6 AA;

Query Match 100.0%; - Score 21; DB 11; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4
 ||||
 Db 2 RGDA 5

RESULT 10

AAR11506
 ID AAR11506 standard; Protein; 6 AA.

XX AAR11506;

XX 12-JUN-1991 (first entry)

XX Cell attachment promoting peptide.

XX Fibrin; aggregation.

XX Synthetic.

XX Key Location/Qualifiers

XX Active-site 2..4

XX US4988621-A.

XX 29-JAN-1991.

XX 10-DEC-1987; 87US-0131130.

XX 10-DEC-1987; 87US-0131130.

XX 24-MAY-1985; 85US-0738078.

XX (JOLL-) LA JOLLA CANCER FOU.

XX Ruoslahti EI, Hayman EG, Pierschbacher MD;

XX WPI; 1991-116404/16.

XX Peptide(s) contg. arginine-glycine-aspartic acid sequence - used
 XX to prevent and reverse cell attachment or to promote cell
 XX aggregation.

XX Disclosure; Page 8; 12pp; English.

XX The peptide, or shorter versions contg. the RGD active site from
 XX fibronectin, can be used to prevent and reverse attachment of cells
 XX to substrates. This can be used in cell prepn., fermentation, cell
 XX line prepn., cell matrix prepn., diagnostics and therapy. The

CC peptide can be used for eg mobilisation of bone marrow cells;
 CC prevention and reversal of attachment of disseminated tumour cells
 CC locally such as in the case of an operation performed in the peri-
 CC toneal cavity, to prevent adhesions and scar formations locally as
 CC in the case of eye operations, for prophylactic inhibition of E. coli
 CC binding to epithelial cells of the urinary tract or intestine,
 CC diagnosis and treatment of E. coli related infections, and
 CC identification of various pathogenic bacterial strains. The
 CC peptide is pref. prepd. by solid phase synthesis.
 CC See also AAR11505

XX Sequence 6 AA;

Query Match 100.0%; Score 21; DB 12; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4
 ||||
 Db 2 RGDA 5

RESULT 11

ABP48385
 ID ABP48385 standard; Peptide; 7 AA.

XX ABP48385;

XX 28-AUG-2002 (first entry)

XX Zinc finger protein related peptide motif SEQ ID NO:289.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.

XX Homo sapiens.

XX Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-USA43438.

XX 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;

XX WPI; 2002-500284/53.

XX New zinc finger protein that binds to target site, useful in studying
 XX gene function and for human therapeutics and plant engineering,
 XX comprises first, second and third zinc fingers, ordered from N- to
 XX C-terminus

XX Example 1; Page 37; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to
 XX a target site, comprising a first (F1), a second (F2), and a third (F3)
 XX zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 XX target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 XX and a third (S3) target subsite. Also described are: (I) a polypeptide
 XX (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 XX (3) designing (M) (I) involves selecting the F1 zinc finger such that
 XX it binds to the S1 target subsite, selecting the F2 zinc finger such
 XX that it binds to the S2 target subsite, and selecting the F3 zinc
 XX finger such that it binds to the S3 target subsite, thus designing (I)
 XX that binds to a target site. (I) is useful for recognition of triplet
 XX target subsites having the nucleotide G in the 5'-most position of the
 XX subsite. (I) is useful in studying gene function, and for human
 XX therapeutics and plant engineering. (I), (II) or (III) is useful in
 XX therapeutic methods to modulate the expression of a target region within

CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determined the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.

CC Sequence 7 AA;
 CC Query Match 100.0%; Score 21; DB 23; Length 7;
 CC Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 Db 1 RGDA 4

RESULT 12

ABP48594
 ID ABP48594 standard; Peptide; 7 AA.

AC ABP48594;
 XX

DT 28-AUG-2002 (first entry)

XX Zinc finger protein related peptide motif SEQ ID NO:670.
 KW Zinc finger protein; ZFP; DNA binding protein; zinc finger.

OS Homo sapiens.
 OS Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US43438.

XX 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;

XX WPI; 2002-500284/53.

XX New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering,
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus

PS Example 1; Page 40; 81pp; English.

CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target sub-site. Also described are: (1) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target sub-site, selecting the F2 zinc finger such that
 CC that it binds to the S2 target sub-site, and selecting the F3 zinc
 CC finger such that it binds to the S3 target sub-site, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target sub-sites having the nucleotide G in the 5'-most position of the
 CC sub-site. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determined the
 CC phenotype and function of gene expression. (I) has improved affinity

CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.

CC Sequence 7 AA;
 CC Query Match 100.0%; Score 21; DB 23; Length 7;
 CC Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 Db 1 RGDA 4

RESULT 13

ABP48597

ID ABP48597 standard; Peptide; 7 AA.

XX AC ABP48597;
 XX

DT 28-AUG-2002 (first entry)

XX Zinc finger protein related peptide motif SEQ ID NO:671.

KW Zinc finger protein; ZFP; DNA binding protein; zinc finger.

OS Homo sapiens.
 OS Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US43438.

XX 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;

XX WPI; 2002-500284/53.

XX New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering,
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus

PS Example 1; Page 40; 81pp; English.

CC The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target sub-site. Also described are: (1) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
 CC that it binds to the S2 target sub-site, and selecting the F3 zinc
 CC finger such that it binds to the S3 target sub-site, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target sub-sites having the nucleotide G in the 5'-most position of the
 CC sub-site. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determined the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given

CC in the exemplification of the present invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 21; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 14

ABP48600
ID ABP48600 standard; Peptide; 7 AA.

XX AC ABP48600;

XX 28-AUG-2002 (first entry)

XX Zinc finger protein related peptide motif SEQ ID NO:672.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.

XX Homo sapiens.

XX Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US43438.

XX 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;

XX WPI; 2002-500284/53.

XX New zinc finger protein that binds to target site, useful in studying
gene function and for human therapeutics and plant engineering,
comprises first, second and third zinc fingers, ordered from N- to
C-terminus

XX Example 1; Page 40; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to
a target site, comprising a first (F1), a second (F2), and a third (F3)
zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
target site comprises, in 3'-5' direction, a first (S1), a second (S2),
and a third (S3) target subsite. Also described are: (1) a polypeptide
(II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
(3) designing (M) (I) involves selecting the F1 zinc finger such that
it binds to the S1 target subsite, selecting the F2 zinc finger such that
it binds to the S2 target subsite, and selecting the F3 zinc
finger such that it binds to the S3 target subsite, thus designing (I)
that binds to a target site. (I) is useful for recognition of triplet
target subsites having the nucleotide G in the 5'-most position of the
subsite. (I) is useful in studying gene function, and for human
therapeutics and plant engineering. (I), (II) or (III) is useful in
therapeutic methods to modulate the expression of a target region within
a subject, in diagnostic methods for sequence specific detection of
target nucleic acid in a sample, and in assays to determine the
phenotype and function of gene expression. (I) has improved affinity
and specificity for their target sequences, as well as enhanced
biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
represent DNA target sequences and zinc finger peptides which are given
in the exemplification of the present invention.

XX Sequence 7 AA;

Query Match 100.0%; Score 21; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Query Match 100.0%; Score 21; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 15

ABP48603
ID ABP48603 standard; Peptide; 7 AA.

XX AC ABP48603;

XX 28-AUG-2002 (first entry)

XX Zinc finger protein related peptide motif SEQ ID NO:673.

XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.

XX Homo sapiens.

XX Synthetic.

XX WO200242459-A2.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US43438.

XX 20-NOV-2000; 2000US-0716637.

XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;

XX WPI; 2002-500284/53.

XX New zinc finger protein that binds to target site, useful in studying
gene function and for human therapeutics and plant engineering,
comprises first, second and third zinc fingers, ordered from N- to
C-terminus

XX Example 1; Page 40; 81pp; English.

XX The present invention describes a zinc finger protein (I) that binds to
a target site, comprising a first (F1), a second (F2), and a third (F3)
zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
target site comprises, in 3'-5' direction, a first (S1), a second (S2),
and a third (S3) target subsite. Also described are: (1) a polypeptide
(II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
(3) designing (M) (I) involves selecting the F1 zinc finger such that
it binds to the S1 target subsite, selecting the F2 zinc finger such
that it binds to the S2 target subsite, and selecting the F3 zinc
finger such that it binds to the S3 target subsite, thus designing (I)
that binds to a target site. (I) is useful for recognition of triplet
target subsites having the nucleotide G in the 5'-most position of the
subsite. (I) is useful in studying gene function, and for human
therapeutics and plant engineering. (I), (II) or (III) is useful in
therapeutic methods to modulate the expression of a target region within
a subject, in diagnostic methods for sequence specific detection of
target nucleic acid in a sample, and in assays to determine the
phenotype and function of gene expression. (I) has improved affinity
and specificity for their target sequences, as well as enhanced
biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
represent DNA target sequences and zinc finger peptides which are given
in the exemplification of the present invention.

XX Sequence 7 AA;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy . 1 RGDA 4
| | | |
Db . 1 RGDA 4

Search completed: May 12, 2003, 15:36:04
Job time : 12.44 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:33:37 ; Search time 3.84 Seconds
(without alignments)
30.649 Million cell updates/sec

Title: US-09-909-348-3

Perfect score: 21

Sequence: 1 RGDA 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	1	US-08-658-130-19
2	21	100.0	4	4	US-09-039-780A-23
3	21	100.0	4	5	PCT-US96-08723-19
4	21	100.0	4	6	5352664-2
5	21	100.0	9	2	US-08-232-081B-3
6	21	100.0	9	4	US-09-644-600-85
7	21	100.0	10	1	US-07-602-847C-1
8	21	100.0	10	1	US-07-602-847C-4
9	21	100.0	10	3	US-09-139-762A-100
10	21	100.0	10	3	US-09-139-762A-113
11	21	100.0	11	1	US-07-609-716-14
12	21	100.0	11	1	US-08-477-509B-35
13	21	100.0	11	3	US-08-482-085B-35
14	21	100.0	11	4	US-08-475-411A-14
15	21	100.0	11	4	US-08-478-029A-14
16	21	100.0	11	4	US-08-444-791A-35
17	21	100.0	14	4	US-08-981-088-4
18	21	100.0	15	1	US-07-602-847C-6
19	21	100.0	15	1	US-07-602-847C-9
20	21	100.0	15	4	US-08-981-088-3
21	21	100.0	16	4	US-09-426-680-2
22	21	100.0	18	1	US-08-176-500-32
23	21	100.0	18	1	US-08-471-052A-32
24	21	100.0	18	1	US-08-189-331-32
25	21	100.0	18	2	US-08-471-939-32
26	21	100.0	18	2	US-08-471-800-32
27	21	100.0	18	2	US-08-471-068-32

28	21	100.0	20	1	US-08-127-351-21	Sequence 21, Appl
29	21	100.0	20	1	US-08-480-367B-21	Sequence 21, Appl
30	21	100.0	20	1	US-08-487-221A-21	Sequence 21, Appl
31	21	100.0	20	1	US-08-480-370-21	Sequence 21, Appl
32	21	100.0	20	6	5196510-29	Patent No. 5196510
33	21	100.0	20	6	5196510-30	Patent No. 5196510
34	21	100.0	20	6	5352664-1	Patent No. 5352664
35	21	100.0	21	1	US-07-609-716-50	Sequence 50, Appl
36	21	100.0	21	1	US-08-842-255-19	Sequence 19, Appl
37	21	100.0	21	4	US-08-475-411A-50	Sequence 50, Appl
38	21	100.0	21	4	US-09-219-849-1	Sequence 1, Appl
39	21	100.0	21	4	US-08-478-029A-50	Sequence 50, Appl
40	21	100.0	23	4	US-08-981-088-2	Sequence 2, Appl
41	21	100.0	23	4	US-08-330-594-8	Sequence 8, Appl
42	21	100.0	23	6	5196510-19	Patent No. 5196510
43	21	100.0	25	1	US-08-127-351-20	Sequence 20, Appl
44	21	100.0	25	1	US-08-480-367B-20	Sequence 20, Appl
45	21	100.0	25	1	US-08-487-221A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-658-130-19
; Sequence 19, Application US/08658130
; Patent No. 5736392
; GENERAL INFORMATION:
; APPLICANT: Hawley-Nelson, Pamela
; APPLICANT: Lan, Jiangling
; APPLICANT: Shih, Pojen
; APPLICANT: Jessee, Joel A.
; APPLICANT: Shifferli, Kevin P.
; TITLE OF INVENTION: Peptide-Enhanced Cationic Lipid
; TITLE OF INVENTION: Transfections
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,130
; FILING DATE: 04-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/477,354
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Caruthers, Jennie M.
; REGISTRATION NUMBER: 34,464
; REFERENCE/DOCKET NUMBER: 32-95A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
US-08-658-130-19

Query Match 100.0%; Score 21; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
||||
Db 1 RGDA 4

RESULT 2

US-09-039-780A-23

; Sequence 23, Application US/09039780A

; Patent No. 6376248

; GENERAL INFORMATION:

; APPLICANT: HAWLEY-NELSON, PAMELA

; LAN, JIANQING

; SHIH, POJEN

; JESSE, JOEL A.

; SCHIFFERLI, KEVIN P.

; GEBEYEHU, GULLILAT

; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

; NUMBER OF SEQUENCES: 120

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GREENLEE, WINNER & SULLIVAN

; STREET: 5370 MANHATTAN CIRCLE, SUITE 201

; CITY: BOULDER

; STATE: CO

; COUNTRY: US

; ZIP: 80303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/039,780A.

; FILING DATE: 16-Mar-1998

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: SULLIVAN, SALLY A.

; REGISTRATION NUMBER: 32,064

; REFERENCE/DOCKET NUMBER: 32-95C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (303)499-8080

; TELEFAX: (303)499-8089

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; SEQUENCE DESCRIPTION: SEQ ID NO: 23:

US-09-039-780A-23

Query Match 100.0%; Score 21; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
||||
Db 1 RGDA 4

RESULT 3

PCT-US96-08723-19

; Sequence 19, Application PC/TUS9608723

; GENERAL INFORMATION:

; APPLICANT: Life Technologies, Inc.

; TITLE OF INVENTION: Peptide-Enhanced Cationic Lipid

; TITLE OF INVENTION: Transfections

Patent No. 5886152
GENERAL INFORMATION:
APPLICANT: NAKATANI, TOMOTUKI
APPLICANT: GOMI, HIDEYUKI
APPLICANT: WIJDENES, JOHN
APPLICANT: NOGUCHI, HIROSHI
TITLE OF INVENTION: HUMANIZED B-B10
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,081B
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 20-3484
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-232-081B-3

Query Match 100.0%; Score 21; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 6
US-09-644-600-85
Sequence 85, Application US/09644600
Patent No. 6451500
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Tanimoto, Hirotochi
TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease
FILE REFERENCE: D6064CIP/D
CURRENT APPLICATION NUMBER: US/09/644,600
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/421,213
PRIOR FILING DATE: 1999-10-20
PRIOR APPLICATION NUMBER: 09/027,337
PRIOR FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 85
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Residues 248-256 of the TAGD-15 protein
US-09-644-600-85

Query Match 100.0%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDA 4
Db 2 RGDA 5
RESULT 7
US-07-602-847C-1
Sequence 1, Application US/07602847C
Patent No. 5227469
GENERAL INFORMATION:
APPLICANT: Lazarus, Robert A.,
APPLICANT: Seymour, Jana L.
TITLE OF INVENTION: NO. 5227469el Platelet Aggregation Inhibitors From The Leec
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/602,847C
FILING DATE: 19901026
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/479,829
FILING DATE: 14-FEB-1990
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Daryl B.
REGISTRATION NUMBER: 32,637
REFERENCE/DOCKET NUMBER: 667
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1249
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-602-847C-1

Query Match 100.0%; Score 21; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 4 RGDA 7

RESULT 8
US-07-602-847C-4
Sequence 4, Application US/07602847C
Patent No. 5227469
GENERAL INFORMATION:
APPLICANT: Lazarus, Robert A.,
APPLICANT: Seymour, Jana L.
TITLE OF INVENTION: NO. 5227469el Platelet Aggregation Inhibitors From The Leec
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/602,847C
FILING DATE: 19901026
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/479,829
FILING DATE: 14-FEB-1990
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Daryl B.
REGISTRATION NUMBER: 32,637
REFERENCE/DOCKET NUMBER: 667
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1249
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-602-847C-4

Query Match 100.0%; Score 21; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 4 RGDA 7

RESULT 9
US-09-139-762A-100
Sequence 100, Application US/09139762A
Patent No. 6013453
GENERAL INFORMATION:
APPLICANT: Choo, Yen
APPLICANT: Klug, Aaron
TITLE OF INVENTION: Improvements in or Relating to
TITLE OF INVENTION: Binding Proteins for Recognition of DNA
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/139,762A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/793,408
FILING DATE: 02-JUN-1997
APPLICATION NUMBER: PCT/GB95/01949
FILING DATE: 17-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9514698.1
FILING DATE: 18-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9422534.9
FILING DATE: 08-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9416880.4
FILING DATE: 20-AUG-1994
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9514698.1
FILING DATE: 18-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9422534.9
FILING DATE: 08-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9416880.4
FILING DATE: 20-AUG-1994
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-139-762A-100

Query Match 100.0%; Score 21; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 10
US-09-139-762A-113
Sequence 113, Application US/09139762A
Patent No. 6013453
GENERAL INFORMATION:
APPLICANT: Choo, Yen
APPLICANT: Klug, Aaron
TITLE OF INVENTION: Improvements in or Relating to
TITLE OF INVENTION: Binding Proteins for Recognition of DNA
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/139,762A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/793,408
FILING DATE: 02-JUN-1997
APPLICATION NUMBER: PCT/GB95/01949
FILING DATE: 17-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9514698.1
FILING DATE: 18-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9422534.9
FILING DATE: 08-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9416880.4
FILING DATE: 20-AUG-1994
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS:

```
;
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-139-762A-113

Query Match 100.0%; Score 21; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4
Db 1 RGDA 4

RESULT 11
US-07-609-716-14
; Sequence 14, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BIR
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-609-716-14

Query Match 100.0%; Score 21; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4
Db 5 RGDA 8

RESULT 12
US-08-477-509B-35
; Sequence 35, Application US/08477509B
; Patent No. 5770697
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W
; APPLICANT: Dorman, Mary A

; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
```

```
;
; TITLE OF INVENTION: No. 5770697el Peptides Comprising Repetitive
; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,509B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,155
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,049
; FILING DATE: 22-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-7/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-477-509B-35

Query Match 100.0%; Score 21; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4
Db 5 RGDA 8

RESULT 13
US-08-482-085B-35
; Sequence 35, Application US/08482085B
; Patent No. 6018030
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Richardson, Charles
; APPLICANT: Chambers, James
; APPLICANT: Causey, Stuart
; APPLICANT: Pollock, Thomas J.
; APPLICANT: Cappello, Joseph
; APPLICANT: Crissman, John W.
; TITLE OF INVENTION: No. 6018030el Peptides Comprising Repetitive
; TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
; NUMBER OF SEQUENCES: 112
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
```

STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,085B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,049
FILING DATE: 22-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,155
FILING DATE: 29-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-6/RFT/MTK
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-482-085B-35

Query Match 100.0%; Score 21; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 5 RGDA 8

RESULT 14
US-08-475-411A-14
Sequence 14, Application US/08475411A
Patent No. 6140072
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Cappello, Joseph
TITLE OF INVENTION: Functional Recombinantly Prepared
TITLE OF INVENTION: Synthetic Protein Polymer
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,411A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/609,716
FILING DATE: 06-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/269,429
FILING DATE: 09-NOV-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-9/RFT/MTK
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-475-411A-14

Query Match 100.0%; Score 21; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 5 RGDA 8

RESULT 15
US-08-478-029A-14
Sequence 14, Application US/08478029A
Patent No. 6184348
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Cappello, Joseph
TITLE OF INVENTION: Functional Recombinantly Prepared
TITLE OF INVENTION: Synthetic Protein Polymer
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,029A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/609,716
FILING DATE: 06-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/269,429
FILING DATE: 09-NOV-1988

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/114,618
;; FILING DATE: 29-OCT-1987
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 06/927,258
;; FILING DATE: 04-NOV-1986
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Trecartin, Richard F.
;; REGISTRATION NUMBER: 21,801
;; REFERENCE/DOCKET NUMBER: A-55186-8/RFT/MTK
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-781-1989
;; TELEFAX: 415-398-3249
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-478-029A-14

Query Match 100.0%; Score 21; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. NO. 39;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 5 RGDA 8

Search completed: May 12, 2003, 15:40:39
Job time : 5.84 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:38:52 ; Search time 7.76 Seconds
(without alignments)
47.436 Million cell updates/sec

Title: US-09-909-348-3

Perfect score: 21

Sequence: 1 RGDA 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	9 US-10-050-692-3	Sequence 3, Appl1
2	21	100.0	4	9 US-10-050-611-1	Sequence 1, Appl1
3	21	100.0	4	9 US-10-050-688-3	Sequence 3, Appl1
4	21	100.0	4	9 US-09-911-569-23	Sequence 23, Appl1
5	21	100.0	4	10 US-09-909-348-3	Sequence 3, Appl1
6	21	100.0	4	10 US-09-904-090-1	Sequence 1, Appl1
7	21	100.0	4	10 US-09-909-122-3	Sequence 3, Appl1
8	21	100.0	7	9 US-09-990-186-289	Sequence 289, App
9	21	100.0	7	9 US-09-990-186-670	Sequence 670, App
10	21	100.0	7	9 US-09-990-186-671	Sequence 671, App
11	21	100.0	7	9 US-09-990-186-672	Sequence 672, App
12	21	100.0	7	9 US-09-990-186-673	Sequence 673, App
13	21	100.0	7	9 US-09-990-186-674	Sequence 674, App
14	21	100.0	7	9 US-09-990-186-675	Sequence 675, App
15	21	100.0	7	9 US-09-990-186-696	Sequence 696, App
16	21	100.0	7	9 US-09-990-186-986	Sequence 986, App
17	21	100.0	7	9 US-09-990-186-971	Sequence 971, App
18	21	100.0	7	9 US-09-990-186-1047	Sequence 1047, App
19	21	100.0	7	9 US-09-990-186-1048	Sequence 1048, App

20	21	100.0	7	9 US-09-990-186-1915	Sequence 1915, Ap
21	21	100.0	7	9 US-09-990-186-1924	Sequence 1924, Ap
22	21	100.0	7	9 US-09-990-186-1925	Sequence 1925, Ap
23	21	100.0	7	9 US-09-990-186-1927	Sequence 1927, Ap
24	21	100.0	7	9 US-09-990-186-1928	Sequence 1928, Ap
25	21	100.0	7	9 US-09-990-186-1961	Sequence 1961, Ap
26	21	100.0	7	9 US-09-990-186-1989	Sequence 1989, Ap
27	21	100.0	7	9 US-09-990-186-1990	Sequence 1990, Ap
28	21	100.0	7	9 US-09-990-186-1996	Sequence 1996, Ap
29	21	100.0	7	9 US-09-990-186-1997	Sequence 1997, Ap
30	21	100.0	7	9 US-09-990-186-2005	Sequence 2005, Ap
31	21	100.0	7	9 US-09-990-186-2007	Sequence 2007, Ap
32	21	100.0	7	9 US-09-990-186-2716	Sequence 2716, Ap
33	21	100.0	7	9 US-09-990-186-2934	Sequence 2934, Ap
34	21	100.0	7	9 US-09-990-186-3169	Sequence 3169, Ap
35	21	100.0	7	9 US-09-990-186-3177	Sequence 3177, Ap
36	21	100.0	7	9 US-09-990-186-3449	Sequence 3449, Ap
37	21	100.0	7	9 US-09-990-186-3702	Sequence 3702, Ap
38	21	100.0	7	9 US-09-990-186-3717	Sequence 3717, Ap
39	21	100.0	7	9 US-09-990-186-3749	Sequence 3749, Ap
40	21	100.0	7	10 US-09-989-789-289	Sequence 289, App
41	21	100.0	7	10 US-09-989-789-670	Sequence 670, App
42	21	100.0	7	10 US-09-989-789-671	Sequence 671, App
43	21	100.0	7	10 US-09-989-789-672	Sequence 672, App
44	21	100.0	7	10 US-09-989-789-673	Sequence 673, App
45	21	100.0	7	10 US-09-989-789-674	Sequence 674, App

ALIGNMENTS

RESULT 1

US-10-050-692-3
; Sequence 3, Application US/10050692
; Publication No. US20020182205A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Simmons, David J.
; APPLICANT: Yang, Jinping
; APPLICANT: Redin, William R.
; TITLE OF INVENTION: STIMULATION OF BONE GROWTH WITH THROMBIN
; FILE REFERENCE: 3033.1002-004
; CURRENT APPLICATION NUMBER: US/10/050,692
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 09/909,122
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/219,300
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human prothrombin
US-10-050-692-3

Query Match 100.0% Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4

Db 1 RGDA 4

RESULT 2

US-10-050-611-1
; Sequence 1, Application US/10050611
; Publication No. US20020187933A1

;; GENERAL INFORMATION:
;; APPLICANT: Carney, Darrell H.
;; TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
;; TITLE OF INVENTION: PEPTIDES
;; FILE REFERENCE: 3033.1000-008
;; CURRENT APPLICATION NUMBER: US/10/050,611
;; PRIOR FILING DATE: 2002-01-16
;; PRIOR APPLICATION NUMBER: 09/904,090
;; PRIOR FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: 60/217,583
;; PRIOR FILING DATE: 2000-07-12
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 4
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: human fragment of thrombin
US-10-050-611-1

Query Match 100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
|||
Db 1 RGDA 4

RESULT 3
US-10-050-688-3
;; Sequence 3, Application US/10050688
;; Publication No. US20020198154A1
;; GENERAL INFORMATION:
;; APPLICANT: Carney, Darrell H.
;; APPLICANT: Crowther, Roger S.
;; APPLICANT: Stierberg, Janet
;; APPLICANT: Bergmann, John
;; TITLE OF INVENTION: STIMULATION OF CARTILAGE GROWTH WITH
;; TITLE OF INVENTION: AGONISTS OF THE NON-PROTEOLYTICALLY ACTIVATED THROMBIN
;; TITLE OF INVENTION: RECEPTOR
;; FILE REFERENCE: 3033.1003-004
;; CURRENT APPLICATION NUMBER: US/10/050,688
;; CURRENT FILING DATE: 2002-01-16
;; PRIOR APPLICATION NUMBER: 09/909,348
;; PRIOR FILING DATE: 2001-07-19
;; PRIOR APPLICATION NUMBER: 60/219,800
;; PRIOR FILING DATE: 2000-07-20
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 4
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: peptide fragment of thrombin
US-10-050-688-3

Query Match 100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
|||
Db 1 RGDA 4

RESULT 4
US-09-911-569-23
;; Sequence 23, Application US/09911569
;; Publication No. US20030069173A1
;; GENERAL INFORMATION:

;; APPLICANT: HAWLEY-NELSON, PAMELA
;; LAN, JIANQING
;; SHIH, POJEN
;; JESSE, JOEL A.
;; SCHIFFERLI, KEVIN P.
;; GEBEYEHU, GULILAT
;; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS
;; NUMBER OF SEQUENCES: 120
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: GREENLEE, WINNER & SULLIVAN
;; STREET: 5370 MANHATTAN CIRCLE, SUITE 201
;; CITY: BOULDER
;; STATE: CO
;; COUNTRY: US
;; ZIP: 80303
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION NUMBER: US/09/911,569
;; FILING DATE: 23-Jul-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 09/039,780
;; FILING DATE: 16-MAR-1998
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SULLIVAN, SALLY A.
;; REGISTRATION NUMBER: 32,064
;; REFERENCE/DOCKET NUMBER: 32-95D
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (303)499-8080
;; TELEFAX: (303)499-8089
;; INFORMATION FOR SEQ ID NO: 23:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-911-569-23

Query Match 100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
|||
Db 1 RGDA 4

RESULT 5
US-09-909-348-3
;; Sequence 3, Application US/09909348
;; Patent No. US20020042373A1
;; GENERAL INFORMATION:
;; APPLICANT: Carney, Darrell H.
;; APPLICANT: Crowther, Roger S.
;; APPLICANT: Stierberg, Janet
;; APPLICANT: Bergmann, John
;; TITLE OF INVENTION: Stimulation Of Cartilage Growth With Agonists
;; FILE REFERENCE: 3033.1003-001
;; CURRENT APPLICATION NUMBER: US/09/909,348
;; CURRENT FILING DATE: 2001-07-19
;; PRIOR APPLICATION NUMBER: US 60/219,800
;; PRIOR FILING DATE: 2000-07-20
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide fragment of Thrombin
US-09-909-348-3

Query Match      100.0%; Score 21; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 6
US-09-904-090-1
; Sequence 1, Application US/09904090
; Patent No. US20020061852A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell
; TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
; FILE REFERENCE: 3033.1000-001
; CURRENT APPLICATION NUMBER: US/09/904,090
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/217,583
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polypeptide, fragment of thrombin
US-09-904-090-1

Query Match      100.0%; Score 21; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 7
US-09-909-122-3
; Sequence 3, Application US/09909122
; Patent No. US20020128202A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Simmons, David J.
; APPLICANT: Yang, Jinping
; APPLICANT: Redin, William R.
; TITLE OF INVENTION: Stimulation Of Bone Growth With Thrombin
; FILE REFERENCE: 3033.1002-001
; CURRENT APPLICATION NUMBER: US/09/909,122
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/219,300
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide fragment of Thrombin
US-09-909-122-3

Query Match      100.0%; Score 21; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 8
US-09-990-186-289
; Sequence 289, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 289
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example 2FP
US-09-990-186-289

Query Match      100.0%; Score 21; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 9
US-09-990-186-670
; Sequence 670, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 670
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example 2FP
US-09-990-186-670

Query Match      100.0%; Score 21; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 10
US-09-990-186-670
; Sequence 670, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 670
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example 2FP
US-09-990-186-670

Query Match      100.0%; Score 21; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 1 RGDA 4

RESULT 10
```



```
US-09-990-186-671
; Sequence 671, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 671
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example zfp
US-09-990-186-671
Query Match 100.0%; Score 21; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.le+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDA 4
| | | |
Db 1 RGDA 4
| | | |
RESULT 11
US-09-990-186-672
; Sequence 672, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 672
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example zfp
US-09-990-186-672
Query Match 100.0%; Score 21; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.le+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDA 4
| | | |
Db 1 RGDA 4
| | | |
RESULT 12
US-09-990-186-673
; Sequence 673, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 673
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example zfp
US-09-990-186-673
Query Match 100.0%; Score 21; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.le+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDA 4
| | | |
Db 1 RGDA 4
| | | |
RESULT 13
US-09-990-186-674
; Sequence 674, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 674
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example zfp
US-09-990-186-674
Query Match 100.0%; Score 21; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.le+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDA 4
| | | |
Db 1 RGDA 4
| | | |
RESULT 14
US-09-990-186-675
; Sequence 675, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 675
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example zfp
US-09-990-186-675
Query Match 100.0%; Score 21; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.le+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RGDA 4
| | | |
Db 1 RGDA 4
| | | |
```

Db 1 RGDA 4

RESULT 15

US-09-990-186-696
; Sequence 696, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; TITLE OF INVENTION: TRIPLETS BY ZINC FINGERS
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 696
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example zfp
US-09-990-186-696

Query Match 100.0%; Score 21; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.le+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RGDA 4

Db 1 RGDA 4

Search completed: May 12, 2003, 15:59:09
Job time : 7.76 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:33:02 ; Search time 4.56 seconds
(without alignments)
84.328 Million cell updates/sec

Title: US-09-909-348-3

Perfect score: 21

Sequence: 1 RGDA 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	19	2	A34467
2	21	100.0	39	2	A36453
3	21	100.0	45	2	G82812
4	21	100.0	49	2	S70093
5	21	100.0	52	2	S19623
6	21	100.0	57	2	E70535
7	21	100.0	68	2	A3217
8	21	100.0	74	2	S2570
9	21	100.0	76	2	I39905
10	21	100.0	79	2	B30870
11	21	100.0	79	2	G85748
12	21	100.0	79	2	E4884
13	21	100.0	80	2	S68677
14	21	100.0	88	2	H82662
15	21	100.0	89	2	I68553
16	21	100.0	90	2	E82562
17	21	100.0	93	2	AH0620
18	21	100.0	95	2	E82696
19	21	100.0	96	2	G84240
20	21	100.0	96	2	D83771
21	21	100.0	97	2	A71034
22	21	100.0	97	2	C75089
23	21	100.0	97	2	E82962
24	21	100.0	98	2	S01566
25	21	100.0	100	2	T30673
26	21	100.0	102	2	E75273
27	21	100.0	103	2	F70976
28	21	100.0	104	2	B72538
29	21	100.0	107	2	F90230

30 21 100.0 108 2 T51207
31 21 100.0 110 2 E97566
32 21 100.0 110 2 AC2787
33 21 100.0 115 2 S14024
34 21 100.0 115 2 C82479
35 21 100.0 116 2 D71832
36 21 100.0 116 2 D64681
37 21 100.0 117 2 B81255
38 21 100.0 121 2 I35719
39 21 100.0 123 2 H75059
40 21 100.0 124 2 D84319
41 21 100.0 124 2 S62816
42 21 100.0 124 2 T03574
43 21 100.0 126 2 C86883
44 21 100.0 126 2 B72621
45 21 100.0 126 2 T37063

hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
ribosomal protein
ribosomal proteina
50S ribosomal prot
phnQ protein - Esc
hypothetical prote
30S ribosomal prot
ribosomal protein
hypothetical prote
50S ribosomal prote
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

A34467

36K microfibril-associated protein - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 08-Jun-1990 #sequence_revision 08-Jun-1990 #text_change 18-Jun-1993

C:Accession: A34467

R:Kobayashi, R.; Tashima, Y.; Masuda, H.; Shozawa, T.; Numata, Y.; Miyauchi, K.; Haya
J. Biol. Chem. 264, 17437-17444, 1989

A:Title: Isolation and characterization of a new 36-kDa microfibril-associated glycop
A:Reference number: A34467; MUID:90008913; PMID:2793866

A:Accession: A34467

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-19 <KOB>

Query Match 100.0%; Score 21; DB 2; Length 19;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4

DB 5 RGDA 8

RESULT 2

A36453

decorsin - leech (Macrobodella decora)

C:Species: Macrobodella decora

C:Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 30-Sep-1993

C:Accession: A36453

R:Seymour, J.L.; Henzel, W.J.; Nevins, B.; Stults, J.T.; Lazarus, R.A.

J. Biol. Chem. 265, 10143-10147, 1990

A:Title: Decorsin. A potent glycoprotein IIB-IIIa antagonist and platelet aggregation
A:Reference number: A36453; MUID:90277628; PMID:2351655

A:Accession: A36453

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-39 <SEY>

Query Match 100.0%; Score 21; DB 2; Length 39;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4

DB 31 RGDA 34

RESULT 3

G82812

hypothetical protein XF0386 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 R;Anonymous: G82812
 A;Title: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
 Nature 406, 151-157, 2000
 A;Reference number: A82515; MUID:20365717; PMID:10910347
 A;Note: for a complete list of authors see reference number A59328 below
 A;Accession: G82812
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-45 <SIM>
 A;Cross-references: GB:AE003890; GB:AE003849; NID:g9105215; PIDN:AAF83196.1; GSPDB:GN001
 A;Experimental source: strain 9a5c
 R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincan, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
 ; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A;Reference number: A59328
 A;Contents: annotation
 C;Genetics:
 A;Gene: XF0386

Query Match 100.0%; Score 21; DB 2; Length 45;
 Best Local Similarity 100.0%; Pred. NO. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 ||||
 Db 19 RGDA 22

RESULT 4
 S70093
 hypothetical protein (orf49) - Amycolatopsis methanolica
 C;Species: Amycolatopsis methanolica
 C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
 C;Accession: S70093
 R;Vrijbloed, J.W.; Jelinkova, M.; Hessels, G.I.; Dijkhuizen, L.
 Mol. Microbiol. 18, 21-31, 1995
 A;Title: Identification of the minimal replicon of plasmid pMEA300 of the methylotrophic
 A;Reference number: S70087; MUID:96154936; PMID:8596458
 A;Accession: S70093
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-49 <VRI>
 A;Cross-references: EMBL:L36679
 C;Genetics:
 A;Start codon: GTG

Query Match 100.0%; Score 21; DB 2; Length 49;
 Best Local Similarity 100.0%; Pred. NO. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 ||||
 Db 23 RGDA 26

RESULT 5
 S19623
 ornatin C - leech (Placobdella ornata)
 C;Species: Placobdella ornata
 C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
 C;Accession: S19623
 R;Mazur, P.; Henzel, W.J.; Seymour, J.L.; Lazarus, R.A.

Eur. J. Biochem. 202, 1073-1082, 1991
 A;Title: Ornatin: potent glycoprotein IIB-IIIa antagonists and platelet aggregation
 A;Reference number: S19566; MUID:92111479; PMID:1765068
 A;Accession: S19623
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-52 <MAZ>

Query Match 100.0%; Score 21; DB 2; Length 52;
 Best Local Similarity 100.0%; Pred. NO. 1.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 ||||
 Db 42 RGDA 45

RESULT 6
 E70535
 hypothetical protein Rv0666 - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C;Accession: E70535
 R;Cole, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: E70535
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-57 <COI>
 A;Cross-references: GB:Z95972; GB:AL123456; NID:g3261790; PIDN:CAB09391.1; PID:e31919
 A;Experimental source: strain H37RV
 C;Genetics:
 A;Gene: Rv0666

Query Match 100.0%; Score 21; DB 2; Length 57;
 Best Local Similarity 100.0%; Pred. NO. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 ||||
 Db 24 RGDA 27

RESULT 7
 AG3217
 hypothetical protein Atu5470 [imported] - Agrobacterium tumefaciens (strain C58, Dupo
 C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C;Accession: AG3217
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; MCCI
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Reference number: AB2577; PMID:11743193
 A;Accession: AG3217
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-68 <KUR>
 A;Cross-references: GB:AE008687; PIDN:AAL46157.1; PID:gl7743927; GSPDB:GN00188
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: Atu5470
 A;Genome: plasmid

Query Match 100.0%; Score 21; DB 2; Length 68;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4

DB 36 RGDA 39

RESULT 8

S62570
608 ribosomal protein l38 - fission yeast (Schizosaccharomyces pombe)
N/Alternate names: protein SPAC30D11.1
C:Species: Schizosaccharomyces pombe
C>Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 11-Jan-2000
C:Accession: S62570; T38587
R:Pearson, D.; Churcher, C.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S62559
A:Accession: S62570
A:Molecule type: DNA
A:Residues: 1-74 <PEA>
A:Cross-references: EMBL:Z67961; NID:gl065887; PIDN:CAA91898.1; PID:gl065899
R:Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z21801
A:Accession: T38587
A:Cross-references: EMBL:Z67961; PIDN:CAA91898.1; GSPDB:GN00066; SPDB:SPAC30D11.12
A:Experimental source: strain 972h; cosmid c30D11
C:Genetics:
A:Gene: rpl38-2; SPAC30D11.12
A:Map position: 1L
A:Introns: 1/3; 64/1
C:Superfamily: rat ribosomal protein l38
C:Keywords: cytosol; protein biosynthesis; ribosome

Query Match 100.0%; Score 21; DB 2; Length 74;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4

DB 17 RGDA 20

RESULT 9

I39905
trp RNA-binding protein - Bacillus pumilus
C:Species: Bacillus pumilus
C>Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 15-Oct-1999
C:Accession: I39905
R:Hoffman, R.J.; Gollnick, P.
J. Bacteriol. 177, 839-842, 1995
A:Title: The trpB gene of Bacillus pumilus encodes a protein with sequence and functional similarity to the trpB gene of Escherichia coli
A:Reference number: I39904; MUID:95138053; PMID:7836324
A:Accession: I39905
A:Cross-references: EMBL:Z67961; NID:gl065887; PIDN:CAA91898.1; PID:gl065899
A:Experimental source: strain 972h; cosmid c30D11
C:Genetics:
A:Gene: trpB

Query Match 100.0%; Score 21; DB 2; Length 76;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4

DB 58 RGDA 61

RESULT 10

B90870
Hypothetical protein Ecs1930 [Imported] - Escherichia coli (strain O157:H7, substrain C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: B90870
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome organization
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90870
A:Cross-references: EMBL:Z67961; NID:gl065887; PIDN:CAA91898.1; PID:gl065899
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: Ecs1930

Query Match 100.0%; Score 21; DB 2; Length 79;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4

DB 5 RGDA 8

RESULT 11

G85748
unknown protein encoded within prophage CP-933R [Imported] - Escherichia coli (strain C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: G85748
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamou, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85748
A:Cross-references: EMBL:Z67961; NID:gl065887; PIDN:CAA91898.1; PID:gl065899
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ydaQ

Query Match 100.0%; Score 21; DB 2; Length 79;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4

DB 5 RGDA 8

RESULT 12

E64884
ydaQ protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: E64884
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: E64884
A:Cross-references: EMBL:Z67961; NID:gl065887; PIDN:CAA91898.1; PID:gl065899
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ydaQ

A:Residues: 1-79 <BLAT>
A:Cross-references: GB:AE000232; GB:U00096; NID:g1787600; PIDN:AACT4428.1; PID:g1787608;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ydaQ

Query Match 100.0%; Score 21; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 5 RGDA 8

RESULT 13
S68677
cytochrome c551 - Chromatium vinosum
C:Species: Chromatium vinosum
C:Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 04-Mar-2000
C:Accession: S68677
R:Samy, B.; de Smet, L.; van Driessche, G.; Meyer, T.E.; Bartsch, R.G.; Cusanovich, M.A.
Eur. J. Biochem. 236, 689-696, 1996
A:Title: A high-potential soluble cytochrome c-551 from the purple phototrophic bacterium
A:Reference number: S68677; MUID:96195682; PMID:8612646
A:Accession: S68677
A:Molecule type: protein
A:Residues: 1-80 <SAM>
A:Experimental source: strain D
C:Superfamily: cytochrome c6; cytochrome c6 homology
C:Keywords: cytochrome; electron transfer; heme; iron; metalloprotein; oxidative phos
F:1-77/Domain: cytochrome c6 homology <CYC>
F:10,13/Binding site: heme (Cys) (covalent) #status predicted
F:14,59/Binding site: heme iron (His, Met) #status predicted

Query Match 100.0%; Score 21; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 33 RGDA 36

RESULT 14
H82662
conserved hypothetical protein XF1562 [Imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: H82662
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82662
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-88 <SIM>
A:Cross-references: GB:AE003986; GB:AE003849; NID:g9106606; PIDN:AAF84371.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Netto, E.; Docena, C.; El-Dorri, H.; Facincan, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovskij-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1562

Query Match 100.0%; Score 21; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 65 RGDA 68

RESULT 15
I68553
cell surface glycoprotein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 23-Jul-1999
C:Accession: I68553
R:Horn, G.T.; Bugawan, T.L.; Long, C.M.; Manos, M.M.; Erlich, H.A.
Hum. Immunol. 21, 249-263, 1998
A:Title: Sequence analysis of HLA class II genes from insulin-dependent diabetic indi
A:Reference number: I54290; MUID:88227495; PMID:3372263
A:Accession: I68553
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <RES>
A:Cross-references: GB:M35000; NID:g291960; PIDN:AAA35774.1; PID:g553265
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein

Query Match 100.0%; Score 21; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
Db 44 RGDA 47

Search completed: May 12, 2003, 15:39:43
Job time: 7.56 secs

GenCore version 5.1.4.p5_4578
Copyright. (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:32:07 ; Search time 2.32 Seconds
(without alignments)
71.511 Million cell updates/sec

Title: US-09-909-348-3

Perfect score: 21

Sequence: 1 RGDA 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	39	1 DECO_MACDE	P17350 macrodella
2	21	100.0	52	1 ORNC_PLAOR	P25512 placodella
3	21	100.0	74	1 R38B_SCHPO	Q09900 schizosacch
4	21	100.0	76	1 WTRB_BACPU	P48064 bacillus pu
5	21	100.0	80	1 C551_CHRVI	P80549 chromatium
6	21	100.0	97	1 RL21_PYRAB	Q9uzp1 pyrococcus
7	21	100.0	97	1 RL21_PYRHO	O74001 pyrococcus
8	21	100.0	98	1 UL19_HCMVA	P16723 human cytom
9	21	100.0	113	1 APGL_HUMAN	Q15772 homo sapien
10	21	100.0	113	1 APGL_MOUSE	Q62407 mus musculus
11	21	100.0	113	1 APGL_RAT	Q63638 rattus norv
12	21	100.0	116	1 RL17_HELPJ	Q9zjt6 helicobacte
13	21	100.0	116	1 RL17_HELPY	P56042 helicobacte
14	21	100.0	124	1 RL17_MYCPN	Q59547 mycoplasma
15	21	100.0	124	1 RS8E_HALNI	Q9hpe9 halobacteri
16	21	100.0	131	1 RL17_THENA	Q9x111 thermotoga
17	21	100.0	133	1 GPEE_BACSU	O06717 bacillus su
18	21	100.0	140	1 CO8B_RAT	P55314 rattus norv
19	21	100.0	141	1 NIKR_METJA	Q57969 methanococc
20	21	100.0	143	1 IR09_HCMVA	P16807 human cytom
21	21	100.0	150	1 FLAG_METVO	O06540 methanococc
22	21	100.0	150	1 MOAE_HAEIN	P45308 haemophilus
23	21	100.0	155	1 RR7_CUSEU	P46292 cuscuta eur
24	21	100.0	157	1 YS10_VIBCH	Q9kuk8 vibrio chol
25	21	100.0	164	1 RL15_HALMA	P12737 haloarcula
26	21	100.0	176	1 YCE7_DROME	Q97067 drosophila
27	21	100.0	177	1 RL6_HALMA	P14135 haloarcula
28	21	100.0	179	1 YF36_PSEAE	Q913h7 pseudomonas
29	21	100.0	185	1 RRF_BUCAI	P57328 buchnera ap
30	21	100.0	190	1 Y2H5_STRCO	P35925 streptomyce
31	21	100.0	192	1 TRPD_ALCSP	P18781 alcaligenes
32	21	100.0	201	1 EFA4_HUMAN	P52798 homo sapien
33	21	100.0	201	1 SODE_ONCVO	Q07449 onchocerca

34 21 100.0 202 1 B3G1_MOUSE
35 21 100.0 203 1 IDI_MYCTU
36 21 100.0 206 1 EFA4_MOUSE
37 21 100.0 206 1 YNAB_BACSU
38 21 100.0 212 1 RB17_HUMAN
39 21 100.0 216 1 ALL7_LEPDS
40 21 100.0 217 1 EVGI_HUMAN
41 21 100.0 224 1 MAUM_PARDE
42 21 100.0 224 1 RADC_VIBCH
43 21 100.0 226 1 AROD_HALNI
44 21 100.0 227 1 UPP_HALNI
45 21 100.0 228 1 LPRH_MYCTU

ALIGNMENTS

RESULT 1

ID DECO_MACDE STANDARD; PRT; 39 AA.
AC P17350;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Decorsin.
OS Macrodella decora (North American leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Macrobdella.
OX NCBI_TaxID=6405;
RN [1]
RP SEQUENCE.
RX MEDLINE=90277628; PubMed=2351655;
RA Seymour J.L., Henzel W.J., Nevins B., Stults J.T., Lazarus R.A.;
RT "Decorsin. A potent glycoprotein IIB-IIIA antagonist and platelet
RT aggregation inhibitor from the leech Macrobdella decora.";
RL J. Biol. Chem. 265:10143-10147(1990).
RN [2]

STRUCTURE BY NMR.

RX MEDLINE=94278502; PubMed=8009227;
RA Krezel A.M., Wagner G., Seymour-Ulmer J., Lazarus R.A.;
RT "Structure of the RGD protein decorsin: conserved motif and distinct
RT function in leech proteins that affect blood clotting.";
RL Science 264:1944-1947(1994).
CC -!- FUNCTION: INHIBITS FIBRINOGEN INTERACTION WITH PLATELET RECEPTORS
CC EXPRESSED ON GLYCOPROTEIN IIB-IIIA COMPLEX. MAY PREVENT BLOOD FROM
CC CLOTTING DURING EITHER FEEDING AND/OR STORAGE OF INGESTED BLOOD.
CC -!- SIMILARITY: HIGH, TO P-ORNATA ORNATINS.
CC -!- SIMILARITY: SOME, TO THE DISINTEGRIN FAMILY.
DR PIR; A36453; A36453.
DR PDB; IDEC; 3I-AUG-94.
KW Blood coagulation; Platelet; Cell adhesion; 3D-structure.
FT DOMAIN 27 38 HIGH AFFINITY BINDING DOMAIN (POTENTIAL).
FT SITE 31 33 CELL ATTACHMENT SITE.
FT VARIANT 1 3 MISSING (IN N-3 ISOFORM).
SQ SEQUENCE 39 AA; 4384 MW; 3A3B35756FB70D36 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGDA 4
| | | |
DB 31 RGDA 34

RESULT 2

ID ORNC_PLAOR STANDARD; PRT; 52 AA.
AC P25512;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ornatin C.

OS Placobdella ornata (Turtle leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 OC Rhynchobdella; Glossiphoniidae; Placobdella.
 OX NCBI_TaxID=6415;
 RN [1]
 RP MEDLINE=92111479; PubMed=1765068;
 RX Mazur P., Henzel W.J., Seymour J.L., Lazarus R.A.;
 RA "ornatins: potent glycoprotein IIB-IIIa antagonists and platelet
 RT aggregation inhibitors from the leech Placobdella ornata.";
 RL Eur. J. Biochem. 202:1073-1082(1991).
 CC -1- FUNCTION: POTENT INHIBITOR OF FIBRINOGEN INTERACTION WITH PLATELET
 CC RECEPTORS EXPRESSED ON GLYCOPROTEIN IIB-IIIa COMPLEX. MAY PREVENT
 CC BLOOD FROM CLOTTING DURING EITHER FEEDING AND/OR STORAGE OF
 CC INGESTED BLOOD.
 CC -1- SIMILARITY: HIGH, TO THE OTHER P. ORNATA ORNATINS, AND TO M. DECORA
 CC DECORSIN.
 CC -1- SIMILARITY: SOME, TO THE DISINTEGRIN FAMILY.
 DR PIR; S19623; S19623.
 DR InterPro: IPR002463; Ornatin.
 DR Pfam: PF02088; Ornatin; 1.
 DR PRINTS; PR01184; ORNATIN.
 KW Blood coagulation; Platelet; Cell adhesion.
 FT SITE 42 44
 FT SITE 42 44
 SQ SEQUENCE 52 AA; 5845 MW; BA55CA7408EF4F09 CRC64;
 Query Match 100.0%; Score 21; DB 1; Length 52;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RGDA 4
 Db 42 RGDA 45

RESULT 3
 R38B SCHPO STANDARD; PRT; 74 AA.
 AC Q09900;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 60S ribosomal protein L38-2.
 GN RPL38B OR RPL38 OR SPAC30D11.12.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkheart G., Aert R., Robben J., Grymonprez B.,
 RA Welljans I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Sipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- MISCELLANEOUS: There are two genes for L38 in S. pombe.
 CC -1- SIMILARITY: BELONGS TO THE L38E FAMILY OF RIBOSOMAL PROTEINS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL; 267961; CAA91898.1;
 DR InterPro: IPR002675; Ribosomal_L38e.
 DR Pfam: PF01781; Ribosomal_L38e; 1.
 DR ProDom: PD010361; Ribosomal_L38e; 1.
 KW Ribosomal protein; Multigene family.
 SQ SEQUENCE 74 AA; 8339 MW; C90D6594DFC11D3 CRC64;
 Query Match 100.0%; Score 21; DB 1; Length 74;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RGDA 4
 Db 17 RGDA 20

RESULT 4
 MTRB_BACPU STANDARD; PRT; 76 AA.
 AC P48064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Transcription attenuation protein mtrB (Tryptophan RNA-binding
 DE attenuator protein) (Trp RNA-binding attenuation protein) (TRAP).
 GN MTRB.
 OS Bacillus pumilus (Bacillus mesentericus).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95138053; PubMed=7836324;
 RA Hoffman R.J., Gollnick P.;
 RT "The mtrB gene of Bacillus pumilus encodes a protein with sequence
 RT and functional homology to the trp RNA-binding attenuation protein
 RT (TRAP) of Bacillus subtilis.";
 RL J. Bacteriol. 177:839-842(1995).
 CC -1- FUNCTION: REQUIRED FOR TRANSCRIPTION ATTENUATION CONTROL IN THE
 CC TRP OPERON. THIS TRANS-ACTING FACTOR SEEMS TO RECOGNIZE A 10 BASES
 CC NUCLEOTIDE SEQUENCE IN THE TRP LEADER TRANSCRIPT CAUSING
 CC TRANSCRIPTION TERMINATION. BINDS THE LEADER RNA ONLY IN PRESENCE
 CC OF L-TRYPTOPHAN.
 CC -1- SUBUNIT: OLIGOMER OF 11 IDENTICAL SUBUNITS ARRANGED IN DOUGHNUT-
 CC LIKE STRUCTURE (BY SIMILARITY).
 CC -1- SIMILARITY: WITH REGA FROM PHAGE T4.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL; L37879; AAA67544.1;
 DR HSSP; Q9X6J6; IQAW.


```

DR InterPro: IPR000824; TrpBP.
DR Pfam: PF02081; TrpBP; 1.
DR PRINTS: PR00687; TRPNAAP.
KW Transcription regulation; RNA-binding.
SQ SEQUENCE 76 AA; 8301 MW; 22184B2351DA151D CRC64;

Query Match 100.0%; Score 21; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4
Db 58 RGDA 61

RESULT 5
C551_CHRVI STANDARD; PRT; 80 AA.
ID C551_CHRVI
AC P80549;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cytochrome c-551 (C551).
OS Chromatium vinosum.
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
OC Allochromatium.
OX NCBI_TaxID=1049;
RN [1]
RP SEQUENCE.
RC STRAIN=D / ATCC 17899 / DSM 180;
RX MEDLINE=96195682; PubMed=9612646;
RA Samyn B., de Smet L., van Driessche G., Meyer T.E., Bartsch R.G.,
RA Cusanovich M.A., van Beeumen J.J.;
RT "A high-potential soluble cytochrome c-551 from the purple
RT phototrophic bacterium Chromatium vinosum is homologous to cytochrome
RL Eur. J. Biochem. 236:689-696(1996).
CC -1- FUNCTION: MONOHEME CYTOCHROME.
DR HSP: P95339; 1A56
DR InterPro: IPR000345; Cyt_c_heme_bind.
DR InterPro: IPR003088; Cyt_Ci.
DR InterPro: IPR002324; Cyt_CiD.
DR Pfam: PF00034; cytochrome_c; 1.
DR PRINTS: PR00606; CYTOCHROME_CiD.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
KW Electron transport; Heme.
FT BINDING 10 10 HEME (COVALENT).
FT BINDING 13 13 HEME (COVALENT).
FT METAL 14 14 IRON (HEME AXIAL LIGAND).
FT METAL 59 59 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE * 80 AA; 8224 MW; EBD30A2815D07F93 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4
Db 33 RGDA 36

RESULT 6
RL21_PVRAB STANDARD; PRT; 97 AA.
ID RL21_PVRAB
AC Q90ZPI;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L21e.
GN RPL21E OR PAB0731.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae;
OC Pyrococcus.

NCBI_TaxID=29292;
[1]
SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L21E FAMILY OF RIBOSOMAL PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: AJ248286; CAB50016.1;
InterPro: IPR001147; Ribosomal_L21e.
Pfam: PF01157; Ribosomal_L21e; 1.
PROSITE: PS01171; RIBOSOMAL_L21E; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 97 AA; 11378 MW; 6CEF3A2DB6A61E40 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RGDA 4
Db 69 RGDA 72

RESULT 7
RL21_PYRHO STANDARD; PRT; 97 AA.
ID RL21_PYRHO
AC O74001;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L21e.
GN RPL21E OR PHS032.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -1- SIMILARITY: BELONGS TO THE L21E FAMILY OF RIBOSOMAL PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: AP000005; BAA30227.1;
InterPro: IPR001147; Ribosomal_L21e.
Pfam: PF01157; Ribosomal_L21e; 1.

```

DR PROSITE: PS01171; RIBOSOMAL_L21E; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 97 AA; 11376 MW; 6D5D29DBFBE0E51 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 97;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 DB 69 RGDA 72

RESULT 8

U119_HCMVA STANDARD; PRT; 98 AA.

AC P16723;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Hypothetical protein U119.
 GN U119.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID-10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88094735; PubMed-2827039;
 RA Beck S., Bartell B.G.;
 RT "Human cytomegalovirus encodes a glycoprotein homologous to MHC
 class-I antigens";
 RL Nature 331:269-272(1988).
 RN [2]

COMPLETE GENOME.

RX MEDLINE-90269039; PubMed-21613119;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 cytomegalovirus strain AD169";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Y00293; ; NOT_ANNOTATED_CDS.
 DR EMBL; X17403; CAA35418.1; -;
 DR PIR; S01566; S01566.
 DR PIR; S09782; S09782.
 KW Hypothetical protein.

SQ SEQUENCE 98 AA; 11280 MW; 7E8A7405611E3F2B CRC64;

Query Match 100.0%; Score 21; DB 1; Length 98;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 DB 95 RGDA 98

RESULT 9

APG1_HUMAN STANDARD; PRT; 113 AA.

AC Q15772;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aortic preferentially expressed protein 1 (APEG-1).
 GN APEG1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96291890; PubMed-8663449;
 RA Hsieh C.-M., Yoshizumi M., Endege W.O., Kho C.-J., Jain M.K.,
 RA Kashiki S., de Los Santos R., Lee W.-S., Perrella M.A., Lee M.-E.;
 RT "APEG-1, a novel gene preferentially expressed in aortic smooth muscle
 cells, is down-regulated by vascular injury";
 RL J. Biol. Chem. 271:17354-17359(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX Tissue-Brain;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -|- FUNCTION: MAY HAVE A ROLE IN REGULATING THE GROWTH AND
 CC DIFFERENTIATION OF ARTERIAL SMOOTH MUSCLE CELLS.
 CC -|- SUBCELLULAR LOCATION: Nuclear.
 CC -|- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN DIFFERENTIATED
 CC ARTERIAL SMOOTH MUSCLE CELLS (ASMC)
 CC -|- DEVELOPMENTAL STAGE: APPEARS TO BE EXPRESSED ONLY IN HIGHLY
 CC DIFFERENTIATED ASMC IN NORMAL VESSEL WALLS AND DOWN-REGULATED IN
 CC DEDIFFERENTIATED ASMC IN VIVO. IN RESPONSE TO VASCULAR INJURIES
 CC ASMC DEDIFFERENTIATE AND CHANGE FROM A QUIESCENT AND CONTRACTILE
 CC PHENOTYPE TO A PROLIFERATIVE AND SYNTHETIC PHENOTYPE. THIS
 CC PROLIFERATION OF VASCULAR SMOOTH MUSCLE CELLS IS ONE OF THE MOST
 CC PROMINENT FEATURES OF ARTERIOSCLEROSIS.
 CC -|- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE DOMAIN.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U57099; AAC50599.1; -;
 DR EMBL; BC006346; AAH06346.1; -;
 DR HSP; P56276; ITLK.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003598; Ig_c2.
 DR Pfam; PF00047; Ig_1.
 DR SMART; SM00408; IGC2; 1.
 KW Immunoglobulin domain; Nuclear protein.
 FT DOMAIN 34 95 IG-LIKE DOMAIN.
 SQ SEQUENCE 113 AA; 12692 MW; 04F367263A1397C5 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 113;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 DB 85 RGDA 88

RESULT 10

APG1_MOUSE STANDARD; PRT; 113 AA.

AC Q62407;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Aortic preferentially expressed protein 1 (APEG-1).
 GN APEG1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=96291890; PubMed=8663449;
RA Hsieh C.-M., Yoshizumi M., Endege W.O., Kho C.-J., Jain M.K.,
RA Kashiki S., de Los Santos R., Lee W.-S., Perrella M.A., Lee M.-E.;
RT "APEG-1, a novel gene preferentially expressed in aortic smooth muscle
RT cells, is down-regulated by vascular injury.";
RL J. Biol. Chem. 271:17354-17359(1996).
CC -1- FUNCTION: MAY HAVE A ROLE IN REGULATING THE GROWTH AND
CC DIFFERENTIATION OF ARTERIAL SMOOTH MUSCLE CELLS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN DIFFERENTIATED
CC ARTERIAL SMOOTH MUSCLE CELLS (ASMC).
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U57098; AAC52666.1; --
DR HSP; P56276; ITLK.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003598; Ig_C2.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00408; IGC2; 1.
DR Immunoglobulin domain; Nuclear protein.
FT DOMAIN 34 95 IG-LIKE DOMAIN.
SQ SEQUENCE 113 AA; 12665 MW; 5F320C5A41C3D870 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
DB 85 RGDA 88

RESULT 11
ID APEGL_RAT STANDARD; PRT; 113 AA.
AC Q63638;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aortic preferentially expressed protein 1 (APEG-1).
GN APEGL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=96291890; PubMed=8663449;
RA Hsieh C.-M., Yoshizumi M., Endege W.O., Kho C.-J., Jain M.K.,
RA Kashiki S., de Los Santos R., Lee W.-S., Perrella M.A., Lee M.-E.;
RT "APEG-1, a novel gene preferentially expressed in aortic smooth muscle
RT cells, is down-regulated by vascular injury.";
RL J. Biol. Chem. 271:17354-17359(1996).
CC -1- FUNCTION: MAY HAVE A ROLE IN REGULATING THE GROWTH AND
CC DIFFERENTIATION OF ARTERIAL SMOOTH MUSCLE CELLS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN DIFFERENTIATED ARTERIAL
CC SMOOTH MUSCLE CELLS (ASMC) IN THE MEDIAL LAYER OF THE AORTA.

```

```

CC WEAKLY DETECTED IN BRAIN AND TESTIS AND TO A LESSER EXTENT IN
CC ORGANS RICH IN STRIATED MUSCLE OR VISCERAL SMOOTH MUSCLE.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U57097; AAC52667.1; --
DR HSP; P56276; ITLK.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003598; Ig_C2.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00408; IGC2; 1.
DR Immunoglobulin domain; Nuclear protein.
FT DOMAIN 34 95 IG-LIKE DOMAIN.
SQ SEQUENCE 113 AA; 12668 MW; B213C366A759A363 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
DB 85 RGDA 88

RESULT 12
ID RL17_HELPJ STANDARD; PRT; 116 AA.
AC Q92JTG;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L17.
GN RPIQ OR JHP1212.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.B., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tunmino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999)
CC -1- SIMILARITY: BELONGS TO THE L17P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001547; AAD06814.1; --
DR InterPro; IPR000456; Ribosomal_L17.
DR Pfam; PF01196; Ribosomal_L17; 1.
DR ProDom; PD004277; Ribosomal_L17; 1.
DR TIGRFAMs; TIGR00059; L17; 1.
DR PROSITE; PS01167; RIBOSOMAL_L17; 1.
DR Ribosomal protein; Complete proteome.
SQ SEQUENCE 116 AA; 13392 MW; EBC77780E2F2F3A1 CRC64;

```

Query Match 100.0%; Score 21; DB 1; Length 116;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 ||||
 DB 104 RGDA 107

RESULT 13

RL17_HELPY STANDARD; PRT; 116 AA.
 AC P56042;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 50S ribosomal protein L17.
 GN RPLQ OR HPI292.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 CC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
 RA Loftus K., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 RT pylori".
 RL Nature 388:539-547(1997).
 CC -1- SIMILARITY: BELONGS TO THE L17P FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF000633; RAD08335.1;
 DR TIGR: HP1292;
 DR InterPro: IPR000456; Ribosomal_L17.
 DR Pfam: PF01196; Ribosomal_L17; 1.
 DR ProDom: PD004277; Ribosomal_L17; 1.
 DR TIGRFAMs: TIGR00059; L17; 1.
 DR PROSITE: PS01167; RIBOSOMAL_L17; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 116 AA; 13364 MW; EBD87890E2F2E4B6 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 116;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 ||||
 DB 104 RGDA 107

RESULT 14

RL17_MYCPN STANDARD; PRT; 124 AA.
 ID RL17_MYCPN
 AC Q59547;
 DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 50S ribosomal protein L17.
 GN RPLQ OR MPN192 OR MP639.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=96177562; PubMed=8604303;
 RA Hilbert H., Himmelfreich R., Plagens H., Herrmann R.;
 RT "Sequence analysis of 56 kb from the genome of the bacterium
 RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and
 RT a cluster of ribosomal protein genes.";
 RL Nucleic Acids Res. 24:628-639(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae.";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SIMILARITY: BELONGS TO THE L17P FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U34795; AAC43689.1;
 DR EMBL: AF000061; AAB96287.1;
 DR InterPro: IPR000456; Ribosomal_L17.
 DR Pfam: PF01196; Ribosomal_L17; 1.
 DR TIGRFAMs: TIGR00059; L17; 1.
 DR PROSITE: PS01167; RIBOSOMAL_L17; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 124 AA; 14245 MW; 3A627DB7EBF8C62E CRC64;

Query Match 100.0%; Score 21; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 ||||
 DB 107 RGDA 110

RESULT 15

RS8E_HALN1 STANDARD; PRT; 124 AA.
 ID RS8E_HALN1
 AC Q9HPE9;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S8e.
 GN RPS8E OR VNG1668G.
 OS Halobacterium sp. (strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,

RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 ET "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -!- SIMILARITY: BELONGS TO THE S8E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE005076; AGI19920.1; -;
 DR InterPro: IPR001047; Ribosomal_S8E.
 DR Pfam: PF01201; Ribosomal_S8e; 1.
 DR PRODOM: PD005658; Ribosomal_S8E; 1.
 DR TIGRFAMs: TIGR00307; S8e; 1.
 DR PROSITE: PS01193; RIBOSOMAL_S8E; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 124 AA; 13515 MW; B7038CF79A83742B CRC64;

Query Match 100.0%; Score 21; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 Db 47 RGDA 50

Search completed: May 12, 2003, 15:36:39
 Job time : 3.32 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:32:37 ; Search time 9.04 Seconds
(without alignments)
91.171 Million cell updates/sec

Title: US-09-909-348-3

Perfect score: 21

Sequence: 1 RGDA 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	45	Q9PGB6	Q9PGB6 xylella fas
2	21	100.0	48	Q9XDV3	Q9XDV3 erythrobaet
3	21	100.0	54	Q8R7H3	Q8R7H3 thermoanaer
4	21	100.0	55	Q8RUZ1	Q8RUZ1 zea mays (m
5	21	100.0	57	Q9N041	Q9N041 macaca fasc
6	21	100.0	57	Q8RUD5	Q8RUD5 zea mays (m
7	21	100.0	57	Q8RUD4	Q8RUD4 zea mays (m
8	21	100.0	57	Q06773	Q06773 mycobacteri
9	21	100.0	58	Q8QS83	Q8QS83 chimpanzee
10	21	100.0	59	Q98LS7	Q98LS7 rhizobium l
11	21	100.0	64	Q8XYQ0	Q8XYQ0 ralstonia s
12	21	100.0	65	Q9N7K3	Q9N7K3 leishmania
13	21	100.0	68	Q8UJK6	Q8UJK6 agrobacteri
14	21	100.0	70	Q8XAV0	Q8XAV0 white spot
15	21	100.0	70	Q8XTW3	Q8XTW3 ralstonia s
16	21	100.0	73	Q8Y128	Q8Y128 ralstonia s

17	21	100.0	75	16	Q8VJ45
18	21	100.0	77	6	Q29171
19	21	100.0	77	16	Q92K10
20	21	100.0	79	16	Q8X8Q7
21	21	100.0	83	17	Q8TK40
22	21	100.0	85	10	Q8W3B8
23	21	100.0	88	16	Q9PD18
24	21	100.0	88	17	Q8ZV78
25	21	100.0	89	5	Q95Y01
26	21	100.0	89	7	Q29783
27	21	100.0	90	16	Q9PAU0
28	21	100.0	91	15	Q9DK41
29	21	100.0	92	9	Q9FZT5
30	21	100.0	93	5	Q9NM57
31	21	100.0	93	10	Q8S2D8
32	21	100.0	93	16	Q8Z7W3
33	21	100.0	95	16	Q9PDS1
34	21	100.0	95	16	Q98FE9
35	21	100.0	96	16	Q9KE84
36	21	100.0	96	17	Q9HR67
37	21	100.0	97	16	Q9HTA8
38	21	100.0	99	2	Q8RM68
39	21	100.0	100	12	Q98239
40	21	100.0	101	2	Q9AJC9
41	21	100.0	102	16	Q9RRN9
42	21	100.0	103	2	Q9LB18
43	21	100.0	103	16	Q06257
44	21	100.0	104	2	Q50073
45	21	100.0	104	17	Q9YBK7

ALIGNMENTS

RESULT 1

Q9PGB6	PRELIMINARY;	PRT;	45 AA.
ID	Q9PGB6		
AC	Q9PGB6;		
DT	01-OCT-2000 (TREMBlrel. 15, Created)		
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)		
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)		
DE	Hypothetical protein Xf0386.		
GN	Xf0386.		
OS	Xylella fastidiosa.		
OC	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;		
OC	Xylella.		
OX	NCBI_TaxID=2371;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-9A5C;		
RX	MEDLINE=20365717; PubMed=10910347;		
RA	Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,		
RA	Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,		
RA	Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,		
RA	Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,		
RA	Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,		
RA	Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,		
RA	Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,		
RA	Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,		
RA	Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,		
RA	Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,		
RA	Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,		
RA	Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,		
RA	Machado M.V., Martins A.M.B.N., Madeira H.M.F., Marino C.L.,		
RA	Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,		
RA	Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,		
RA	Moore D.H., Nagai M.A., Nascimento A.L.F.O., Netto L.E.S.,		
RA	Nhanl A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,		
RA	de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,		
RA	Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,		
RA	Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,		
RA	de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,		
RA	da Silva A.C.R., da Silva F.R., da Silva W.A. Jr.,		

da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubato M.H.,
 RA Vallada H., Van Sluys M.A., Verjowski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
 RL Nature 406:151-159(2000).
 DR EMBL; AF003890; AAF83196.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 45 AA; 5163 MW; B58C9AEC98090C8A CRC64;

Query Match 100.0%; Score 21; DB 16; Length 45;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 Db 19 RGDA 22

RESULT 2
 Q9XDV3 PRELIMINARY; PRT; 48 AA.
 AC Q9XDV3;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ORF Q.
 OS Erythrobacter sp. MBIC3960.
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
 OC Erythrobacter.
 OX NCBI_TaxID=94771;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MBIC3960;
 RA Hamada T.;
 RT "Nucleotide sequences of genes coding for photosynthetic reaction
 RT centers and light-harvesting proteins of Erythrobacter litoralis and
 RT related aerobic photosynthetic bacteria.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB027515; BAA78669.1; -;
 DR InterPro; IPR001552; Acyl-CoA_dh.
 DR PROSITE; PS00073; ACYL_COA_DH_2; UNKNOWN.1.
 SQ SEQUENCE 48 AA; 4980 MW; D663EAD05EA8079B CRC64;

Query Match 100.0%; Score 21; DB 2; Length 48;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 Db 27 RGDA 30

RESULT 3
 Q8R7H3 PRELIMINARY; PRT; 54 AA.
 AC Q8R7H3;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein TTE2436.
 GN TTE2436.
 OS Thermoanaerobacter tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MB4T / JCM11007;
 RX MEDLINE-21992816; PubMed-11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;

"A complete sequence of *T. tengcongensis* genome.";
 RL Genome Res. 12:689-700(2002).
 DR EMBL; AEO13185; AAM25571.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 54 AA; 6252 MW; 0A9C818C07DD905B CRC64;

Query Match 100.0%; Score 21; DB 16; Length 54;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 Db 32 RGDA 35

RESULT 4
 Q8RUZ1 PRELIMINARY; PRT; 55 AA.
 AC Q8RUZ1;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Acetyl-CoA C-acyltransferase-like protein (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoidae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VARIOUS STRAINS;
 RA Ching A.S., Caldwell K.S., Jung M., Dolan M., Smith O.S., Tingey S.,
 RA Morgante M., Rafalski J.A.;
 RT "SNP frequency, haplotype structure and linkage disequilibrium in
 RT elite maize inbred lines.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF498463; AAM14479.1; -;
 DR EMBL; AF498463; AAM14485.1; -;
 DR EMBL; AF498472; AAM14488.1; -;
 DR EMBL; AF498477; AAM14493.1; -;
 DR EMBL; AF498482; AAM14498.1; -;
 DR EMBL; AF498485; AAM14501.1; -;
 DR EMBL; AF498486; AAM14502.1; -;
 KW Acyltransferase; Transferase.
 FT NON_TER 1
 SQ SEQUENCE 55 AA; 5959 MW; 5C09DAC7224451D0 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 55;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
 Db 31 RGDA 34

RESULT 5
 Q9N041 PRELIMINARY; PRT; 57 AA.
 AC Q9N041;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE Unnamed protein product.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;

RT "isolation of full-length cDNA clones from macaque brain cDNA
libraries."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046091; BAB01673.1; -
SQ SEQUENCE 57 AA; 6250 MW; 300DE046A4A4897A9 CRC64;

Query Match 100.0%; Score 21; DB 6; Length 57;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
DB 10 RGDA 13

RESULT 6
Q8RUD5 ID Q8RUD5 PRELIMINARY; PRT; 57 AA.

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Acetyl-CoA C-acyltransferase-like protein (Fragment).
OS Zea mays (Maize).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-VARIOUS STRAINS;
RA Ching A.S., Caldwell K.S., Jung M., Dolan M., Smith O.S., Tingey S.,
RA Morgante M., Rafalski J.A.;
RT "SNP frequency, haplotype structure and linkage disequilibrium in
elite maize inbred lines";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF498457; AAM14473.1; -
DR EMBL; AF498458; AAM14474.1; -
DR EMBL; AF498459; AAM14475.1; -
DR EMBL; AF498460; AAM14476.1; -
DR EMBL; AF498461; AAM14477.1; -
DR EMBL; AF498462; AAM14478.1; -
DR EMBL; AF498463; AAM14479.1; -
DR EMBL; AF498464; AAM14480.1; -
DR EMBL; AF498465; AAM14481.1; -
DR EMBL; AF498466; AAM14482.1; -
DR EMBL; AF498467; AAM14483.1; -
DR EMBL; AF498468; AAM14484.1; -
DR EMBL; AF498470; AAM14486.1; -
DR EMBL; AF498471; AAM14487.1; -
DR EMBL; AF498473; AAM14489.1; -
DR EMBL; AF498475; AAM14491.1; -
DR EMBL; AF498478; AAM14496.1; -
DR EMBL; AF498481; AAM14497.1; -
DR EMBL; AF498483; AAM14499.1; -
DR EMBL; AF498484; AAM14500.1; -
DR EMBL; AF498487; AAM14503.1; -
KW Acyltransferase; Transferase.

FT NON_TER 1
SQ SEQUENCE 57 AA; 6203 MW; DC4596C27A4451A8 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 57;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
DB 33 RGDA 36

RESULT 7
Q8RUD4 ID Q8RUD4 PRELIMINARY; PRT; 57 AA.

AC Q8RUD4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Acetyl-CoA C-acyltransferase-like protein (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-CV. IVANA, CV. D71-4HT, AND CV. H60;
RA Ching A.S., Caldwell K.S., Jung M., Dolan M., Smith O.S., Tingey S.,
RA Morgante M., Rafalski J.A.;
RT "SNP frequency, haplotype structure and linkage disequilibrium in
elite maize inbred lines";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF498474; AAM14490.1; -
DR EMBL; AF498476; AAM14492.1; -
DR EMBL; AF498479; AAM14495.1; -
KW Acyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 57 AA; 6185 MW; DC4596C7684451A8 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 57;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
DB 33 RGDA 36

RESULT 8

O06773 ID O06773 PRELIMINARY; PRT; 57 AA.
AC O06773;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein Rv0666.
DE Rv0666 OR MTC1376.10C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;
RX MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., Mclean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence";
RL Nature 393:537-544(1998).
DR EMBL; Z95972; CAB09391.1; -
DR Tuberculinist; Rv0666; -
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 57 AA; 5849 MW; 62858455BD7D0F2E CRC64;
Query Match 100.0%; Score 21; DB 16; Length 57;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGDA 4
DB 11


```

Db      24  RGDA 27

RESULT 9
Q8XQ00;
ID      Q8XQ00; PRELIMINARY; PRT; 64 AA.
AC      Q8XQ00;
DT      01-MAR-2002 (TREMBlrel. 20, Created)
DT      01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT      01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE      Hypothetical protein RSC1708.
GN      RSC1708 OR RS02894.
OS      Ralstonia solanacearum (Pseudomonas solanacearum).
OC      Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC      Ralstonia.
OX      NCBI_TaxID=305;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=GM11000;
RX      MEDLINE=21681879; PubMed=11823852;
RA      Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA      Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA      Chander M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA      Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA      Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA      Weissbach J., Boucher C.A.;
RT      "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL      Nature 415:497-502(2002).
DR      EMBL; AL646066; CAD15410.1;
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 64 AA; 7210 MW; F35FAEF5E609609 CRC64;

Query Match      100.0%; Score 21; DB 16; Length 64;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  RGDA 4
      ||||
Db      60  RGDA 63

RESULT 10
Q98LS7;
ID      Q98LS7; PRELIMINARY; PRT; 59 AA.
AC      Q98LS7;
DT      01-OCT-2001 (TREMBlrel. 18, Created)
DT      01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT      01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE      Hypothetical protein ms10897.
GN      MS10897.
OS      Rhizobium loti (Mesorhizobium loti).
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Phyllobacteriaceae; Mesorhizobium.
OX      NCBI_TaxID=381;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MAFF303099;
RX      MEDLINE=21082930; PubMed=11214968;
RA      Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA      Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA      Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA      MOChizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA      Takeuchi C., Yamada M., Tabata S.;
RT      "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT      Mesorhizobium loti.";
RL      DNA Res. 7:331-338(2000).
DR      EMBL; AP002996; BAB48386.1;
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 59 AA; 6059 MW; 4EE7FEF3940E6633 CRC64;

Query Match      100.0%; Score 21; DB 16; Length 59;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  RGDA 4
      ||||
Db      36  RGDA 39

RESULT 11
Q8XQ00;
ID      Q8XQ00; PRELIMINARY; PRT; 68 AA.
AC      Q8XQ00;
DT      01-JUN-2002 (TREMBlrel. 21, Created)
DT      01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE      chimpanzee cytomegalovirus.
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC      Betaherpesvirinae; Cytomegalovirus.
OX      NCBI_TaxID=188763;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MAFF303099;
RX      MEDLINE=21082930; PubMed=11214968;
RA      Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA      Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA      Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA      MOChizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA      Takeuchi C., Yamada M., Tabata S.;
RT      "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT      Mesorhizobium loti.";
RL      DNA Res. 7:331-338(2000).
DR      EMBL; AP002996; BAB48386.1;
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 59 AA; 6059 MW; 4EE7FEF3940E6633 CRC64;

Query Match      100.0%; Score 21; DB 16; Length 59;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  RGDA 4
      ||||
Db      36  RGDA 39

RESULT 12
Q9N7K3;
ID      Q9N7K3; PRELIMINARY; PRT; 65 AA.
AC      Q9N7K3;
DT      01-OCT-2000 (TREMBlrel. 15, Created)
DT      01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT      01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE      Probable hypothetical 32.4 kDa protein (fragment).
GN      LM28.210.
OS      Leishmania major.
OC      Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX      NCBI_TaxID=5664;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FRIEDLIN;
RA      Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AL390935; CAC00934.1;
FT      NON_TER 1
FT      NON_TER 65
SQ      SEQUENCE 65 AA; 7500 MW; 1169F7CE2F5C12 CRC64;

Query Match      100.0%; Score 21; DB 5; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  RGDA 4
      ||||
Db      28  RGDA 31

RESULT 13
Q8UJK6;
ID      Q8UJK6; PRELIMINARY; PRT; 68 AA.
AC      Q8UJK6;
DT      01-JUN-2002 (TREMBlrel. 21, Created)
DT      01-JUN-2002 (TREMBlrel. 21, Last sequence update)

```

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Hypothetical protein AtU5470.
 GN ATU5470 OR AGR PAT_693.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Plasmid AT.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Rhizobium.
 OK NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura Y., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houniel K., Gordon J., Vaubin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE008968; AAL46157.1; -;
 DR EMBL; AE007916; AAK90845.1; -;
 SQ SEQUENCE 68 AA; 8005 MW; 5CABE406D75F93A8 CRC64;
 Query Match 100.0%; Score 21; DB 16; Length 68;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 RGDA 4
 Db 36 RGDA 39
 RESULT 14
 Q8VAV0 PRELIMINARY; PRT; 70 AA.
 ID Q8VAV0
 AC Q8VAV0;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Wsv279 (WSSV334).
 OS White spot syndrome virus (WSSV).
 OC Viruses; unclassified viruses.
 OK NCBI_TaxID=92652;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21548311; PubMed=11689662;
 RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
 RT "Complete genome sequence of the shrimp white spot bacilliform
 virus.";
 RL J. Virol. 75:11811-11820(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN-TAIWAN;
 RX MEDLINE=20517548; PubMed=11062040;
 RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
 RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
 RT "Identification and characterization of a shrimp white spot syndrome
 virus (WSSV) gene that encodes a novel chimeric polypeptide of
 cellular-type thymidine kinase and thymidylate kinase.";
 RL Virology 277:100-110(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TAIWAN;
 RX MEDLINE=21844071; PubMed=11853398;
 RA Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
 RA Lo C.F., Kou G.H.;
 RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white
 spot syndrome virus and characterization of the motif important for
 targeting VP35 to the nuclei of transfected insect cells.";
 RL Virology 293:44-53(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TAIWAN;
 RA Lo C.-F., Kou G.-H.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF332093; AAL33282.1; -;
 DR EMBL; AF440570; AAL89202.1; -;
 SQ SEQUENCE 70 AA; 7417 MW; 6C31737FF210FB62 CRC64;
 Query Match 100.0%; Score 21; DB 12; Length 70;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 RGDA 4
 Db 30 RGDA 33
 RESULT 15
 Q8XTW3 PRELIMINARY; PRT; 70 AA.
 ID Q8XTW3
 AC Q8XTW3;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein RSC3433.
 GN RSC3433 OR RS01813.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OK NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GMI1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salancoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weissbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).
 DR EMBL; AL646075; CAD16930.1; -;
 SQ SEQUENCE 70 AA; 7019 MW; D0F4CCEA5D6F49CB CRC64;
 Query Match 100.0%; Score 21; DB 16; Length 70;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 RGDA 4
 Db 16 RGDA 19

Search completed: May 12, 2003, 15:38:43
Job time : 12.04 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:31:22 ; Search time 71.5 Seconds
(without alignments)
46.591 Million cell updates/sec

Title: US-09-909-348-5

Perfect score: 141

Sequence: 1 AGTRYKPDGKRCACGDSGGPFV 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	121	85.8	23	20	AAW83414
2	121	85.8	23	21	AA112893
3	121	85.8	23	22	AA112893
4	121	85.8	23	23	AA112893
5	121	85.8	23	23	AA112893
6	121	85.8	23	23	AA112893
7	121	85.8	116	20	AAW99115
8	121	85.8	259	18	AAW11545
9	121	85.8	295	16	AAW74775
10	121	85.8	295	16	AAW74776

11	121	85.8	295	16	AAW74777	Mutant thrombin E2
12	121	85.8	295	16	AAW74778	Mutant thrombin E2
13	121	85.8	295	16	AAW74779	Mutant thrombin E2
14	121	85.8	295	16	AAW74780	Mutant thrombin E2
15	121	85.8	295	16	AAW76033	Mutant thrombin R2
16	121	85.8	295	16	AAW76034	Mutant thrombin R2
17	121	85.8	295	16	AAW76035	Mutant thrombin R2
18	121	85.8	295	16	AAW76036	Mutant thrombin R2
19	121	85.8	295	16	AAW76037	Mutant thrombin W5
20	121	85.8	295	16	AAW76038	Mutant thrombin W5
21	121	85.8	295	16	AAW76039	Mutant thrombin W5
22	121	85.8	295	16	AAW76040	Mutant thrombin W5
23	121	85.8	295	18	AAW22892	Human mature throm
24	121	85.8	295	21	AAW08833	Amino acid sequenc
25	121	85.8	308	20	AAW99109	Human prothrombin
26	121	85.8	376	14	AAW41797	CD4/Thrombin fusio
27	121	85.8	376	20	AAW42789	Human CD4/thrombin
28	121	85.8	376	23	AAU10703	Human CD4-thrombin
29	121	85.8	579	14	AAW35763	Prothrombin (PT)
30	121	85.8	579	18	AAW11546	Human prothrombin
31	121	85.8	579	18	AAW11544	Human prothrombin
32	121	85.8	579	20	AAW99108	Human prothrombin
33	121	85.8	615	14	AAW38741	Human prothrombin
34	121	85.8	615	17	AAW96216	Human prothrombin
35	121	85.8	615	17	AAW90377	Human prothrombin
36	121	85.8	622	18	AAW11543	Human prothrombin
37	121	85.8	622	20	AAW49566	Platelet membrane
38	114	80.9	23	23	AAW78376	Thrombin peptide d
39	114	80.9	111	20	AAW99113	Bovine zeta 2 pret
40	114	80.9	308	20	AAW99107	Bovine prothrombin
41	114	80.9	582	20	AAW99106	Bovine prothrombin
42	81	57.4	15	18	AAW10750	Thrombin B chain a
43	81	57.4	15	22	AAW83282	Modified RGD pepti
44	81	57.4	15	22	AAW83283	Modified RGD pepti
45	80	56.7	259	16	AAW76041	Mutant thrombin se

ALIGNMENTS

RESULT 1
AAW83414
ID AAW83414 standard; peptide; 23 AA.
XX
AC AAW83414;
XX
DT 26-FEB-1999 (first entry)
XX
DE Cell growth/adhesion promoting peptide #1.
XX
KW Cell growth; adhesion; promotion; medical treatment; injury;
KW Biotissue; bone reinforcement; nerve regeneration; HMP resin.
XX
OS Synthetic.
XX
PN JF10316581-A.
XX
PD 02-DEC-1998.
XX
PF 15-MAY-1997; 97JP-0140885.
XX
PR 15-MAY-1997; 97JP-0140885.
XX
PA (KURS) KURARAY CO LTD.
XX
DR WPI; 1999-076400/07.
XX
PT Material for medical treatment comprises new peptide - used for
PT covering injuries, promoting adhesion of bio-tissues, bone
XX reinforcing and nerve regeneration
PS Claim 1; Page 12; 14pp; Japanese.
XX

CC The present invention describes a material for medical treatment which
 CC comprises one or more peptides of the formula XADGSLMPROOY, or their
 CC salts, immobilised on a substrate: where X = H, CH3CO or CH3COLys;
 CC A = Ser or Thr; D = Ile, Val or Leu; E = Lys or Arg; G = Ile, Val or
 CC Leu; J = Gly or Ala; L = Ile, Val or Leu; M = Gly or Ala; Q = Gly, Ala
 CC or Gly-Lys-Lys-Gly; Y = OH or NH2. Also described is an agent for cell
 CC growth promotion and/or cell adhesion promotion containing the above
 CC peptide or its salt as the active component. The peptide and its salt
 CC can be used for covering injuries, promoting adhesion of biotissues,
 CC bone reinforcing and nerve regeneration. The present sequence represents
 CC a specifically claimed peptide of the present invention.

XX Sequence 23 AA;

Query Match 85.8%; Score 121; DB 20; Length 23;
 Best Local Similarity 100.0%; Pred. NO. 1.6e-07;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
 Db 3 YKPDEGKRGDACEGDSGGPFV 23

RESULT 2

AAB12893
 ID AAB12893 standard; peptide; 23 AA.

XX AC AAB12893;

XX DT 02-NOV-2000 (first entry)

XX DE Nerve tissue regenerative peptide SEQ ID #8.

XX KW Nerve regeneration; nerve cell proliferation; axon extension; treatment;
 KW central nervous system disorder; peripheral nervous system disorder;
 KW spinal disorder; head injury; cerebrovascular disorder.

XX OS Synthetic.

XX PN JP2000143531-A.

XX PD 23-MAY-2000.

XX PF 11-AUG-1999; 99JP-0227108.

XX PR 09-SEP-1998; 98JP-0270498.

XX PA (KURS) KURARAY CO LTD.

XX PA (NISH/) NISHIMURA Y.

XX PA (SUZU/) SUZUKI Y.

XX PA (TANI/) TANIHARA M.

XX DR WPI; 2000-415772/36.

XX PT New nerve regeneration material

XX PS Claim 2; Page 5; 17pp; Japanese.

XX CC This invention relates to a new nerve regenerative material which
 CC contains a peptide immobilised to a base which consists of a
 CC polysaccharide gel such as alginic acid. Sequences AAB12886-B12899
 CC represent examples of the peptides used in the nerve regeneration
 CC material. The peptide containing material causes nerve cell
 CC proliferation and also causes axonal extension. The material can be used
 CC for the treatment of central or peripheral nervous system disorders,
 CC spinal disorders, head injury or cerebrovascular disorders.

XX Sequence 23 AA;

Query Match 85.8%; Score 121; DB 21; Length 23;
 Best Local Similarity 100.0%; Pred. NO. 1.6e-07;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
 Db 3 YKPDEGKRGDACEGDSGGPFV 23

RESULT 3

AAB70363
 ID AAB70363 standard; peptide; 23 AA.

XX AC AAB70363;

XX DT 02-MAY-2001 (first entry)

XX DE Human thrombin receptor binding domain peptide SEQ ID NO:8.

XX KW Neutrophil cell chemotactic; wound healing; inflammation; vulnery;
 KW antiinflammatory.

XX OS Homo sapiens.

XX PN US6184342-B1.

XX PD 06-FEB-2001.

XX PF 28-OCT-1994; 94US-0330594.

XX PR 28-OCT-1994; 94US-0330594.

XX PA (CHRY-) CHRYSALIS BIOTECHNOLOGY INC.

XX PI Carney DH, Ramakrishnan S;

XX DR WPI; 2001-202003/20.

XX PT New synthetic neutrophil cell chemotactic peptides, useful for
 PT generating antibodies for modulating neutrophil chemotaxis in immune
 PT response and wound healing

XX PS Example 2; Column 6; 15pp; English.

XX CC The present invention describes a synthetic peptide (I) which is a
 CC neutrophil cell chemotactic agent. (I) has vulnery and
 CC antiinflammatory activities. (I) is useful as a potent neutrophil cell
 CC chemotactic agent and for generating antibodies against the peptides,
 CC which are useful for modulating neutrophil recruitment to a wound site
 CC for enhancing or inhibiting inflammation and early effects of wound
 CC healing. Neutrophil response to (I) is specific, since monocytes and
 CC fibroblasts do not show any expression of the receptor to which (I)
 CC binds. The present sequence represents a human thrombin receptor binding
 CC domain peptide which is used in an example from the present invention.

XX SQ Sequence 23 AA;

Query Match 85.8%; Score 121; DB 22; Length 23;
 Best Local Similarity 100.0%; Pred. NO. 1.6e-07;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
 Db 3 YKPDEGKRGDACEGDSGGPFV 23

RESULT 4

AAE22563
 ID AAE22563 standard; peptide; 23 AA.

XX AC AAE22563;

XX DT 26-JUL-2002 (first entry)

XX DE Human thrombin high affinity receptor binding domain.

XX KW Human; proteolytically activated receptor for thrombin; neutrophil;

KW chemotactic agent; PART; inflammation; wound healing; chemotaxis;
 KW immune response; vulnery; thrombin; receptor binding domain.
 OS Homo sapiens.

XX US2002032314-A1.
 XX 14-MAR-2002.

XX 05-FEB-2001; 2001US-0777328.
 XX 28-OCT-1994; 94US-0330594.

XX (CHRY-) CHRYSALIS BIOTECHNOLOGY INC.
 XX Carney DH, Ramakrishnan S;
 XX WPI; 2002-371207/40.

XX New synthetic peptide neutrophil cell chemotactic agents, useful for
 PT stimulating or modulating neutrophil cell chemotactic migration,
 PT particularly for modulating neutrophil recruitment during immune
 PT response or in wound healing -

XX Example 2; Page 3; 15pp; English.

XX The present invention relates to novel synthetic peptides and antibodies
 CC which are potent chemotactic agents for neutrophils. The peptides of the
 CC invention mimic the activity and role of the cleavage fragment of the
 CC proteolytically activated receptor for thrombin (PART). They are useful
 CC for stimulating or modulating neutrophil cell chemotactic migration or
 CC for generating an antibody. In particular, the peptides of the invention
 CC are useful for modulating neutrophil recruitment to a wound site for
 CC enhancing or inhibiting inflammation and early effects in wound healing.
 CC They are also useful for modulated neutrophil chemotaxis in immune
 CC response. The present sequence is high affinity receptor binding
 CC domain of human thrombin. This peptide is used in the exemplification
 CC of the invention.

XX Sequence 23 AA;

Query Match 85.8%; Score 121; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGRGDACEGDSGGPFV 25
 DB 3 YKPDEGRGDACEGDSGGPFV 23

RESULT 5

AAE20159
 ID AAE20159 standard; peptide; 23 AA.

XX AAE20159;

XX 18-JUN-2002 (first entry)

XX Human thrombin peptide derivative #2.

XX Cartilage growth; cartilage repair; arthritic joint; traumatic injury;
 KW non-proteolytically activated thrombin receptor; NPAP; chondrocyte;
 KW therapy; implantation; thrombin peptide; human.

OS Homo sapiens.

XX WO200207748-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US22668.

XX 20-JUL-2000; 2000US-219800P.

XX PA
 XX PI
 XX DR
 XX DR
 XX PT
 XX PT
 XX PS
 XX PS
 XX CC
 XX CC
 XX CC
 XX CC
 XX CC
 XX CC
 XX CC
 XX CC
 XX SQ

(TEXA) UNIV TEXAS SYSTEM.

Carney DH, Crowther RS, Stiernberg J, Bergmann J;
 WPI; 2002-268953/31.

Stimulating growth and repair of cartilage, useful for treating e.g.
 arthritis, by local administration of an agonist of non-proteolytically
 activated thrombin receptor -

Claim 12; Page 25; 28pp; English.

The invention relates to a method of stimulating growth and repair of
 cartilage. The method involves administering to the site, an agonist
 of non-proteolytically activated thrombin receptor (NPAP). The method
 is used in human or veterinary medicine for the treatment of arthritic
 joints and damage/loss of cartilage caused by traumatic injury. Also
 chondrocytes may be cultured in presence of NPAP agonist to provide
 cells for implantation at sites requiring growth/repair of cartilage.
 The present sequence is human thrombin peptide derivative which serves
 as a NPAP agonist.

XX Sequence 23 AA;

Query Match 85.8%; Score 121; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGRGDACEGDSGGPFV 25
 DB 3 YKPDEGRGDACEGDSGGPFV 23

RESULT 6

AAAM50858

ID AAAM50858 standard; Peptide; 23 AA.

XX AAAM50858;

XX 01-MAY-2002 (first entry)

XX Thrombin-derived peptide used to promote cardiac tissue repair.

XX Thrombin; revascularisation; vascular occlusion; tissue repair;
 KW vulnery; vasotropic; cardiant; angiogenesis; restenosis;
 KW therapy; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 10..13

FT Peptide /note= "thrombin receptor binding domain"

FT Peptide 12..23

FT Peptide /note= "serine esterase conserved sequence"

XX WO200204008-A2.

XX 17-JAN-2002.

XX 12-JUL-2001; 2001WO-US21944.

XX 12-JUL-2000; 2000US-217583P.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Carney DH;

XX WPI; 2002-179665/23.

XX Promoting cardiac tissue repair, stimulating revascularisation,
 PT stimulating vascular endothelial cell proliferation, and inhibiting
 PT vascular occlusion by using angiogenic thrombin derivative peptide

CC of prothrombin. The half-life may be reduced to less than 10 minutes
 CC or the mutant prothrombin may have an extended half-life of more than
 CC 1 hour, making it useful as an anticoagulant and to inhibit side-
 CC effects of anti-coagulant treatment. They are converted to inactive
 CC thrombin and are able to compete with native, active thrombin for
 CC binding to receptors. The present sequence represents the thrombin
 CC mutant which is derived by trypsin cleavage of a specifically
 CC claimed human prothrombin mutant in which Asp at position 419 is
 CC changed to Asn. The thrombin Asn99 mutant was found to have only
 CC 0.24% of the activity of wild-type thrombin on a chromogenic
 CC substrate.
 CC (Note: This sequence does not appear in the specification and has
 CC been produced by modifying the wild-type sequence of human
 CC prothrombin which appears in figure 1).

XX SQ Sequence 259 AA;

Query Match 85.8%; Score 121; DB 18; Length 259;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
 |||||
 Db 190 YKPDGKRGDACEGDSGGPFV 210

RESULT 9

AAR74775
 ID AAR74775 standard; Protein: 295 AA.

XX AC AAR74775;

XX DT 04-NOV-1995 (first entry)

XX DE Wild-type thrombin.

XX KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;
 XX KW anticoagulant; protein engineering; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX FT Protein 37..295
 XX FT /note= "mature protein"

XX PN WO9513385-A.

XX PD 18-MAY-1995.

XX PF 14-NOV-1994; 94WO-US13104.

XX PR 10-JUN-1994; 94US-0258038.

XX PR 12-NOV-1993; 93US-0152657.

XX PA (GILE-) GILEAD SCI.

XX PI Gibbs CS, Leung LLK, Tsiang M;

XX DR WPI; 1995-194103/25.

XX DR N-PSDB; AAQ92455.

XX PT Thrombin derives with segregated pro- and anticoagulant activities
 XX PT useful for treating thrombotic disorders but also diagnosis,
 XX PT treatment of tumours, etc.

XX PS Disclosure; Fig 1; 78pp; English.

XX CC The sequence represents wild-type (reference) thrombin. Mutants
 XX CC of this sequence (AAR74776-80 and AAR76033-41) have at least 80%
 XX CC homology with thrombin, and are capable of protein-C activation
 XX CC without significant fibrinogen clotting activity, and vice versa
 XX CC (specifically have a ratio of protein-C activity to fibrinogen
 XX CC clotting activity of less than 0.5 or greater than 2 compared to

CC thrombin). The mutant thrombin sequences, produced in recombinant
 CC cell culture or by in vitro methods, and are used to treat
 CC thrombotic conditions, particularly during cardiac bypass surgery
 CC and in cases of septic shock.

XX SQ Sequence 295 AA;

Query Match 85.8%; Score 121; DB 16; Length 295;
 Best Local Similarity 100.0%; Pred. No. 1.6e-06;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
 |||||
 Db 226 YKPDGKRGDACEGDSGGPFV 246

RESULT 10

AAR74776

ID AAR74776 standard; Protein: 295 AA.

XX AC AAR74776;

XX DT 04-NOV-1995 (first entry)

XX DE Mutant thrombin K52A, R233A.

XX KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;
 XX KW anticoagulant; protein engineering; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX FT Misc-difference 88 /note= "Iys in wild-type"
 XX FT Misc-difference 269 /note= "Arg in wild-type"
 XX FT Protein 37..295 /note= "mature protein"

XX PN WO9513385-A.

XX PD 18-MAY-1995.

XX PF 14-NOV-1994; 94WO-US13104.

XX PR 10-JUN-1994; 94US-0258038.

XX PR 12-NOV-1993; 93US-0152657.

XX PA (GILE-) GILEAD SCI.

XX PI Gibbs CS, Leung LLK, Tsiang M;

XX DR WPI; 1995-194103/25.

XX PT Thrombin derives with segregated pro- and anticoagulant activities
 XX PT useful for treating thrombotic disorders but also diagnosis,
 XX PT treatment of tumours, etc.

XX PS Claim 22; Page 63/3; 78pp; English.

XX CC The mutant thrombin sequence, generated by oligonucleotide-directed
 XX CC mutagenesis, has at least 80% homology with thrombin, and is
 XX CC capable of protein-C activation without significant fibrinogen
 XX CC clotting activity, and vice versa (specifically, it has a ratio
 XX CC of protein-C activity to fibrinogen clotting activity of less than
 XX CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin
 XX CC is produced in recombinant cell culture or by in vitro methods,
 XX CC and is used to treat thrombotic conditions, particularly during
 XX CC cardiac bypass surgery and in cases of septic shock.

XX SQ Sequence 295 AA;

Query Match 85.8%; Score 121; DB 16; Length 295;

Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
|||||
Db 226 YKPDEGKRGDACEGDSGGPFV 246

RESULT 11

AAAR74777
ID AAR74777 standard; Protein; 295 AA.

AC AAR74777;
XX

DT 04-NOV-1995 (first entry)
XX

DE Mutant thrombin E229D.
XX

KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;
KW antioagulant; protein engineering; ss.
XX

OS Homo sapiens.
XX

XX Key Location/Qualifiers
FH Misc-difference 265

FT /note= "Glu in wild-type"
FT Protein
FT 37..295

FT /note= "mature protein"
XX

PN WO9513385-A.
XX

PD 18-MAY-1995.
XX

PF 14-NOV-1994; 94WO-US13104.
XX

PR 10-JUN-1994; 94US-0258038.
PR 12-NOV-1993; 93US-0152657.

XX (GILE-) GILEAD SCI.
XX

PI Gibbs CS, Leung LLK, Tsiang M;
DR WPI; 1995-194103/25.

XX Thrombin derivs with segregated pro- and antioagulant activities
PT useful for treating thrombotic disorders but also diagnosis,
PT treatment of tumours, etc.

XX Claim 22; Page 63/3; 78pp; English.
PS

XX The mutant thrombin sequence, generated by oligonucleotide-directed
CC mutagenesis, has at least 80% homology with thrombin, and is
CC capable of protein-C activation without significant fibrinogen
CC clotting activity, and vice versa (specifically, it has a ratio
CC of protein-C activity to fibrinogen clotting activity of less than
CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin
CC is produced in recombinant cell culture or by in vitro methods,
CC and is used to treat thrombotic conditions, particularly during
CC cardiac bypass surgery and in cases of septic shock.

XX Query Match 85.8%; Score 121; DB 16; Length 295;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
|||||
Db 226 YKPDEGKRGDACEGDSGGPFV 246

RESULT 12

AAAR74778

ID AAR74778 standard; Protein; 295 AA.

AC AAR74778;
XX

DT 04-NOV-1995 (first entry)
XX

DE Mutant thrombin E229F.
XX

KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;
KW antioagulant; protein engineering; ss.
XX

OS Homo sapiens.
XX

XX Key Location/Qualifiers
FH Misc-difference 265

FT /note= "Glu in wild-type"
FT Protein
FT 37..295

FT /note= "mature protein"
XX

PN WO9513385-A.
XX

PD 18-MAY-1995.
XX

PF 14-NOV-1994; 94WO-US13104.
XX

PR 10-JUN-1994; 94US-0258038.
PR 12-NOV-1993; 93US-0152657.

XX (GILE-) GILEAD SCI.
XX

PI Gibbs CS, Leung LLK, Tsiang M;
DR WPI; 1995-194103/25.

XX Thrombin derivs with segregated pro- and antioagulant activities
PT useful for treating thrombotic disorders but also diagnosis,
PT treatment of tumours, etc.

XX Claim 22; Page 63/3; 78pp; English.
PS

XX The mutant thrombin sequence, generated by oligonucleotide-directed
CC mutagenesis, has at least 80% homology with thrombin, and is
CC capable of protein-C activation without significant fibrinogen
CC clotting activity, and vice versa (specifically, it has a ratio
CC of protein-C activity to fibrinogen clotting activity of less than
CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin
CC is produced in recombinant cell culture or by in vitro methods,
CC and is used to treat thrombotic conditions, particularly during
CC cardiac bypass surgery and in cases of septic shock.

XX Query Match 85.8%; Score 121; DB 16; Length 295;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
|||||
Db 226 YKPDEGKRGDACEGDSGGPFV 246

RESULT 13

AAAR74779

ID AAR74779 standard; Protein; 295 AA.

AC AAR74779;
XX

DT 04-NOV-1995 (first entry)
XX

DE Mutant thrombin E229S.
XX

KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;
KW antioagulant; protein engineering; ss.
XX

```

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 265
XX FT /note= "Glu in wild-type"
XX FT Protein 37..295
XX FT /note= "mature protein"
XX PN WO9513385-A.
XX PD 18-MAY-1995.
XX PF 14-NOV-1994; 94WO-US13104.
XX PR 10-JUN-1994; 94US-0258038.
XX PR 12-NOV-1993; 93US-0152657.
XX PA (GILE-) GILEAD SCI.
XX PI Gibbs CS, Leung LK, Tsiang M;
XX PI WPI; 1995-194103/25.
XX DR Thrombin derivs with segregated pro- and anticoagulant activities
XX DR useful for treating thrombotic disorders but also diagnosis,
XX DR treatment of tumours, etc.
XX PS Claim 22; Page 63/3; 78pp; English.
XX CC The mutant thrombin sequence, generated by oligonucleotide-directed
XX CC mutagenesis, has at least 80% homology with thrombin, and is
XX CC capable of protein-C activation without significant fibrinogen
XX CC clotting activity, and vice versa (specifically, it has a ratio
XX CC of protein-C activity to fibrinogen clotting activity of less than
XX CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin
XX CC is produced in recombinant cell culture or by in vitro methods,
XX CC and is used to treat thrombotic conditions, particularly during
XX CC cardiac bypass surgery and in cases of septic shock.
XX SQ Sequence 295 AA;
XX Query Match 85.8%; Score 121; DB 16; Length 295;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 5 YKPDGKRGDACEGDSGGPFV 25
XX DB |||||
XX RESULT 14
XX AAR74780
XX ID AAR74780 standard; Protein; 295 AA.
XX AC AAR74780;
XX DT 04-NOV-1995 (first entry)
XX DE Mutant thrombin E229W.
XX KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;
XX KW anticoagulant; protein engineering; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 265
XX FT /note= "Glu in wild-type"
XX FT Protein 37..295
XX FT /note= "mature protein"
XX PN WO9513385-A.

```

```

XX PD 18-MAY-1995.
XX PF 14-NOV-1994; 94WO-US13104.
XX PR 10-JUN-1994; 94US-0258038.
XX PR 12-NOV-1993; 93US-0152657.
XX PA (GILE-) GILEAD SCI.
XX PI Gibbs CS, Leung LK, Tsiang M;
XX PI WPI; 1995-194103/25.
XX DR Thrombin derivs with segregated pro- and anticoagulant activities
XX DR useful for treating thrombotic disorders but also diagnosis,
XX DR treatment of tumours, etc.
XX PS Claim 22; Page 63/3; 78pp; English.
XX CC The mutant thrombin sequence, generated by oligonucleotide-directed
XX CC mutagenesis, has at least 80% homology with thrombin, and is
XX CC capable of protein-C activation without significant fibrinogen
XX CC clotting activity, and vice versa (specifically, it has a ratio
XX CC of protein-C activity to fibrinogen clotting activity of less than
XX CC 0.5 or greater than 2 compared to thrombin). The mutant thrombin
XX CC is produced in recombinant cell culture or by in vitro methods,
XX CC and is used to treat thrombotic conditions, particularly during
XX CC cardiac bypass surgery and in cases of septic shock.
XX SQ Sequence 295 AA;
XX Query Match 85.8%; Score 121; DB 16; Length 295;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-06;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 5 YKPDGKRGDACEGDSGGPFV 25
XX DB |||||
XX RESULT 15
XX AAR76033
XX ID AAR76033 standard; Protein; 295 AA.
XX AC AAR76033;
XX DT 04-NOV-1995 (first entry)
XX DE Mutant thrombin E229Y.
XX KW Thrombin; oligonucleotide-directed mutagenesis; procoagulant;
XX KW anticoagulant; protein engineering; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 265
XX FT /note= "Glu in wild-type"
XX FT Protein 37..295
XX FT /note= "mature protein"
XX PN WO9513385-A.
XX PD 18-MAY-1995.
XX PF 14-NOV-1994; 94WO-US13104.
XX PR 10-JUN-1994; 94US-0258038.
XX PR 12-NOV-1993; 93US-0152657.
XX PA (GILE-) GILEAD SCI.
XX PN WO9513385-A.

```

Tue May 13 09:55:20 2003

PI	Gibbs CS, Leung LLK, Tsiang M;
XX	
DR	WPI: 1995-194103/25.
XX	
PT	Thrombin derivs with segregated pro- and anticoagulant activities
XX	useful for treating thrombotic disorders but also diagnosis,
PT	treatment of tumours, etc.
XX	
PS	Claim 22; Page 63/3; 78pp; English.
XX	
CC	The mutant thrombin sequence, generated by oligonucleotide-directed
CC	mutagenesis, has at least 80% homology with thrombin, and is
CC	capable of protein-C activation without significant fibrinogen
CC	clotting activity, and vice versa (specifically, it has a ratio
CC	of protein-C activity to fibrinogen clotting activity of less than
CC	0.5 or greater than 2 compared to thrombin). The mutant thrombin
CC	is produced in recombinant cell culture or by <i>in vitro</i> methods,
CC	and is used to treat thrombotic conditions, particularly during
CC	cardiac bypass surgery and in cases of septic shock.
XX	
XX	Sequence 295 AA;
SQ	

```

Query Match      85.8%   Score 121;   DB 15;   Length 295;
Best Local Similarity 100.0%;   Pred. No. 1.6e-06;
Matches 21;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```

Search completed: May 12, 2003, 15:36:06
Job time : 72.5 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:33:37 ; Search time 24 Seconds
(without alignments)
30.649 Million cell updates/sec

Title: US-09-909-348-5

Perfect score: 141

Sequence: 1 ACTRYKDEKRGDACEGDSGGPFV 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	85.8	23	4	US-08-981-088-2
2	121	85.8	23	4	US-08-330-594-8
3	121	85.8	259	4	US-08-944-483-52
4	121	85.8	295	3	US-08-338-368-2
5	121	85.8	376	2	US-08-538-269-10
6	121	85.8	376	4	US-09-410-882-10
7	121	85.8	579	1	US-08-295-411-4
8	121	85.8	579	2	US-08-955-471-4
9	121	85.8	579	3	US-09-117-708-14
10	121	85.8	579	5	PCT-US92-10242-4
11	121	85.8	615	1	US-07-998-972A-3
12	121	85.8	615	1	US-08-463-953-3
13	121	85.8	615	1	US-08-462-261-3
14	121	85.8	615	5	PCT-US92-11357-3
15	121	85.8	622	3	US-08-952-967-8
16	117	83.0	42	1	US-08-293-778-11
17	104	73.8	20	6	5352664-1
18	81	57.4	15	4	US-08-981-088-3
19	71	50.4	14	4	US-08-981-088-4
20	71	50.4	250	4	US-08-944-483-51
21	71	50.4	261	6	5270178-5
22	71	50.4	261	6	5270178-19
23	71	50.4	261	6	5270178-20
24	71	50.4	261	6	5270178-21
25	71	50.4	262	1	US-07-720-189-1
26	71	50.4	409	4	US-09-065-872-2
27	71	50.4	409	4	US-09-667-570A-2

28	71	50.4	410	4	US-09-065-872-1	Sequence 1, Appl1
29	71	50.4	410	4	US-09-667-570A-1	Sequence 1, Appl1
30	71	50.4	419	1	US-08-295-411-1	Sequence 1, Appl1
31	71	50.4	419	2	US-08-955-471-1	Sequence 1, Appl1
32	71	50.4	419	4	US-09-667-570A-3	Sequence 3, Appl1
33	71	50.4	419	5	PCT-US92-10242-1	Sequence 1, Appl1
34	71	50.4	460	2	US-08-756-506-2	Sequence 2, Appl1
35	71	50.4	460	2	US-08-756-506-4	Sequence 4, Appl1
36	71	50.4	460	6	5270178-13	Patent No. 5270178
37	71	50.4	460	6	5270178-14	Patent No. 5270178
38	71	50.4	460	6	5270178-15	Patent No. 5270178
39	71	50.4	460	6	5270178-16	Patent No. 5270178
40	71	50.4	461	6	5225537-2	Patent No. 5225537
41	71	50.4	461	6	5270178-2	Patent No. 5270178
42	71	50.4	461	6	5270178-17	Patent No. 5270178
43	71	50.4	461	6	5270178-18	Patent No. 5270178
44	71	50.4	461	6	5460953-3	Patent No. 5460953
45	69	48.9	12	6	5352664-3	Patent No. 5352664

ALIGNMENTS

RESULT 1
US-08-981-088-2
; Sequence 2, Application US/08981088
; Patent No. 6146824
; GENERAL INFORMATION:
; APPLICANT: BAR-SHAVIT, RACHEL
; TITLE OF INVENTION: ANTI-METASTATIC AND ANTI-ANGIOGENIC
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,088
; FILING DATE: 27-JAN-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: OLIFF, JAMES A 27,075
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: JAO 40455
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)-836-6400
; TELEFAX: (703)-836-2787
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ; HYPOTHETICAL: NO
US-08-981-088-2

Query Match 85.8% Score 121; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YKPDGKRGDACEGDSGGPFV 25
Db 3 YKPDGKRGDACEGDSGGPFV 23

RESULT 2

US-08-330-594-8
; Sequence 8, Application US/08330594C
; Patent No. 6184342
; GENERAL INFORMATION:
; APPLICANT: CAREY, DARRELL H.
; APPLICANT: RAMAKRISHNAN, SHYAM
; TITLE OF INVENTION: SYNTHETIC PEPTIDE NEUTROPHIL CELL CHEMOTACTIC AGENTS
; FILE REFERENCE: CHBP:002
; CURRENT APPLICATION NUMBER: US/08/330,594C
; CURRENT FILING DATE: 1994-10-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-08-330-594-8

Query Match 85.8%; Score 121; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
|||||

DB 3 YKPDEGKRGDACEGDSGGPFV 23
|||||

RESULT 3

US-08-944-483-52
; Sequence 52, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:

; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 259 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-52

Query Match 85.8%; Score 121; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
|||||

DB 190 YKPDEGKRGDACEGDSGGPFV 210
|||||

RESULT 4

US-08-338-368-2
; Sequence 2, Application US/08338368
; Patent No. 6110721
; GENERAL INFORMATION:
; APPLICANT: GIBBS, CRAIG S.
; APPLICANT: LEUNG, LAWRENCE L.K.
; APPLICANT: TSIANG, MANUEL
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND COAGULATION
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GILEAD SCIENCES, INC.
; STREET: 353 LAKESIDE DRIVE
; CITY: FOSTER CITY
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94404

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/338,368
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,038
; FILING DATE: 10-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: HENSLEY, MAX D.
; REGISTRATION NUMBER: 27,043
; REFERENCE/DOCKET NUMBER: 190.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-574-3000
; TELEFAX: 415-573-4899

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-338-368-2

Query Match 85.8%; Score 121; DB 3; Length 295;
Best Local Similarity 100.0%; Pred. No. 8.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
|||||

DB 226 YKPDEGKRGDACEGDSGGPFV 246
|||||

```
RESULT 5
US-08-558-269-10
; Sequence 10, Application US/08558269
; Patent No. 5961973
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,269
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,800
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CRI-001CP2
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-558-269-10

Query Match 85.88; Score 121; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.le-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
|||||
Db 307 YKPDGKRGDACEGDSGGPFV 327

RESULT 6
US-09-410-882-10
; Sequence 10, Application US/09410882
; Patent No. 6287561
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; TITLE OF INVENTION: PATHOGEN-TARGETED BIOCATALYSTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/410,882
; FILING DATE: 22-AUG-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,989
; FILING DATE: 18-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI263.0C1
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-410-882-10
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/410,882
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,269
; FILING DATE: 13-NOV-1995
; APPLICATION NUMBER: US 07/847,800
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CRI-001CP2
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-410-882-10

Query Match 85.88; Score 121; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.le-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
|||||
Db 307 YKPDGKRGDACEGDSGGPFV 327

RESULT 7
US-08-295-411-4
; Sequence 4, Application US/08295411
; Patent No. 5679639
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Mesters, Rolf M.
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 No. 5679639th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/295,411
; FILING DATE: 22-AUG-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,989
; FILING DATE: 18-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI263.0C1
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
```

```
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 579 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 1..320
;; OTHER INFORMATION: /note= "Prothrombin Light Chain"
;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 321..579
;; OTHER INFORMATION: /note= "Prothrombin Heavy Chain"
US-08-295-411-4
Query Match 85.8%; Score 121; DB 1; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 YKPDEGKRGDAGEGDSGGPFV 25
Db 510 YKPDEGKRGDAGEGDSGGPFV 530
RESULT 8
US-08-955-471-4
; Sequence 4, Application US/08955471
; Patent No. 5968751
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Mesters, Rolf M.
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 No. 596875th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/955,471
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/295,411
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI263.0C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 1..320
;; OTHER INFORMATION: /note= "Prothrombin Light Chain"
;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 321..579
;; OTHER INFORMATION: /note= "Prothrombin Heavy Chain"
US-08-955-471-4
Query Match 85.8%; Score 121; DB 2; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 YKPDEGKRGDAGEGDSGGPFV 25
Db 510 YKPDEGKRGDAGEGDSGGPFV 530
RESULT 9
US-09-117-708-14
; Sequence 14, Application US/09117708
; Patent No. 6060300
; GENERAL INFORMATION:
; APPLICANT: RADITSHC, Martin; FRIEDRICH, Thomas;
; APPLICANT: BOLLSCHWEILER, Claus; SCHMIDT, Martin;
; APPLICANT: HOFFKEN, Hans Wolfgang; SCHWEDEN, Juergen;
; APPLICANT: and RUEBSAMEN, Klaus
; TITLE OF INVENTION: Thrombin mutants as antidote for
; TITLE OF INVENTION: thrombin inhibitors
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kell & Weinkeuf
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM AT-compatible, Pentium processor
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect version 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/117,708
; FILING DATE: 04-AUG-1998
; CLASSIFICATION: 424
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-117-708-14
Query Match 85.8%; Score 121; DB 3; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 YKPDEGKRGDAGEGDSGGPFV 25
Db 510 YKPDEGKRGDAGEGDSGGPFV 530
RESULT 10
PCT-US92-10242-4
; Sequence 4, Application PC/TUS9210242
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Mesters, Rolf
```

```

; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10242
; FILING DATE: 19921118
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,989
; FILING DATE: 18-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRO472P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..320
; OTHER INFORMATION: /note= "Prothrombin Light Chain"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 321..579
; OTHER INFORMATION: /note= "Prothrombin Heavy Chain"
; PCT-US92-10242-4
;
; Query Match 85.8%; Score 121; DB 5; Length 579;
; Best Local Similarity 100.0%; Pred. No. 1.7e-09;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 5 YKPDGKRGDACEGDSGGPFV 25
; Db 510 YKPDGKRGDACEGDSGGPFV 530
;
; RESULT 11
; US-07-998-972A-3
; Sequence 3, Application US/07998972A
; Patent No. 5476777
; GENERAL INFORMATION:
; APPLICANT: Holly, Richard D.
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: METHODS FOR PRODUCING THROMBIN
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Street Tower,
; CITY: Twentieth Floor
; STATE: San Francisco
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,953
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/860,701
; FILING DATE: 31-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/816,281
; FILING DATE: 31-DEC-1991
;
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10242
; FILING DATE: 19921118
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,989
; FILING DATE: 18-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRO472P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 579 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..320
; OTHER INFORMATION: /note= "Prothrombin Light Chain"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 321..579
; OTHER INFORMATION: /note= "Prothrombin Heavy Chain"
; PCT-US92-10242-4
;
; Query Match 85.8%; Score 121; DB 1; Length 615;
; Best Local Similarity 100.0%; Pred. No. 1.8e-09;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 5 YKPDGKRGDACEGDSGGPFV 25
; Db 546 YKPDGKRGDACEGDSGGPFV 566
;
; RESULT 12
; US-08-463-953-3
; Sequence 3, Application US/08463953
; Patent No. 5502034
; GENERAL INFORMATION:
; APPLICANT: Holly, Richard D.
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: METHODS FOR PRODUCING THROMBIN
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Street Tower,
; CITY: Twentieth Floor
; STATE: San Francisco
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,953
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/860,701
; FILING DATE: 31-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/816,281
; FILING DATE: 31-DEC-1991

```



```

; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-12-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-953-3

Query Match      85.8%; Score 121; DB 1; Length 615;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy  5 YKPDGKRGDACEGDSGGPFV 25
Db  546 YKPDGKRGDACEGDSGGPFV 566

RESULT 14
PCT-US92-11357-3
; Sequence 3, Application PC/TUS9211357
; GENERAL INFORMATION:
; APPLICANT: Holly, Richard D.
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: METHODS FOR PRODUCING THROMBIN
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Street Tower,
; STREET: Twentieth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/11357
; FILING DATE: 19921230
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/860,701
; FILING DATE: 31-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/816,281
; FILING DATE: 31-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-12-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-11357-3

Query Match      85.8%; Score 121; DB 5; Length 615;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy  5 YKPDGKRGDACEGDSGGPFV 25
Db  546 YKPDGKRGDACEGDSGGPFV 566

RESULT 15
US-08-952-967-8
; Sequence 8, Application US/08952967
; Patent No. 6086871
; GENERAL INFORMATION:
; APPLICANT: Fischer, Bernhard
; APPLICANT: Schlokot, Uwe
; APPLICANT: Mitterer, Artur
; APPLICANT: Falkner, Falko-Guenter

```

APPLICANT: Eibl, Johann
TITLE OF INVENTION: PROTHROMBIN DERIVATIVES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/952,967
FILING DATE: 26-JAN-1998
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AT96/00105
FILING DATE: 12-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AT A 1006/95
FILING DATE: 13-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Isacson, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 065691/0127
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 622 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-952-967-8

Query Match 85.8%; Score 121; DB 3; Length 622;
Best Local Similarity 100.0%; Pred. NO. 1.8e-09;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 YKPECKRGDACEGSGGPFV 25

|||||

Db 553 YKPECKRGDACEGSGGPFV 573

Search completed: May 12, 2003, 15:40:40
Job time : 24 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:38:52 ; Search time 48.5 Seconds
(without alignments)
47.436 Million cell updates/sec

Title: US-09-909-348-5

Perfect score: 141

Sequence: 1 AGTRYKPDGKRGDACEGDSGGPFV 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141	100.0	25	9 US-10-050-692-5	Sequence 5, Appli
2	141	100.0	25	10 US-09-909-348-5	Sequence 5, Appli
3	141	100.0	25	10 US-09-909-122-5	Sequence 5, Appli
4	121	85.8	23	9 US-10-050-611-3	Sequence 3, Appli
5	121	85.8	23	9 US-10-050-611-4	Sequence 4, Appli
6	121	85.8	23	9 US-10-050-688-5	Sequence 5, Appli
7	121	85.8	23	9 US-10-050-688-6	Sequence 6, Appli
8	121	85.8	23	10 US-09-777-328-8	Sequence 8, Appli
9	121	85.8	23	10 US-09-904-090-3	Sequence 3, Appli
10	118	83.7	23	9 US-10-050-692-6	Sequence 6, Appli
11	114	80.9	250	9 US-09-898-837A-45	Sequence 45, Appli
12	114	80.9	251	9 US-09-898-837A-41	Sequence 41, Appli
13	71	50.4	419	9 US-10-182-263-1	Sequence 1, Appli
14	71	50.4	419	9 US-10-182-263-3	Sequence 3, Appli
15	71	50.4	419	9 US-10-182-263-4	Sequence 4, Appli
16	71	50.4	419	9 US-10-182-263-5	Sequence 5, Appli
17	71	50.4	419	9 US-10-182-263-6	Sequence 6, Appli
18	71	50.4	419	9 US-09-978-917A-4	Sequence 4, Appli
19	71	50.4	461	9 US-10-182-263-2	Sequence 2, Appli

20	71	50.4	451	9 US-09-978-917A-2	Sequence 2, Appli
21	69	48.9	12	9 US-10-050-611-2	Sequence 2, Appli
22	69	48.9	12	10 US-09-904-090-2	Sequence 2, Appli
23	68.5	48.6	260	9 US-10-366-035-2	Sequence 2, Appli
24	68.5	48.6	290	9 US-10-041-006A-7	Sequence 7, Appli
25	68.5	48.6	290	9 US-10-028-072-222	Sequence 222, App
26	68.5	48.6	290	9 US-10-121-049-222	Sequence 222, App
27	68.5	48.6	290	9 US-10-123-904-222	Sequence 222, App
28	68.5	48.6	290	9 US-10-140-470-222	Sequence 222, App
29	68.5	48.6	290	9 US-10-175-746-222	Sequence 222, App
30	68.5	48.6	290	9 US-10-176-918-222	Sequence 222, App
31	68.5	48.6	290	9 US-10-176-921-222	Sequence 222, App
32	68.5	48.6	290	9 US-10-137-865-222	Sequence 222, App
33	68.5	48.6	290	9 US-10-140-474-222	Sequence 222, App
34	68.5	48.6	290	9 US-10-142-431-222	Sequence 222, App
35	68.5	48.6	290	9 US-10-143-114-222	Sequence 222, App
36	68.5	48.6	290	9 US-10-140-002-222	Sequence 222, App
37	68.5	48.6	290	9 US-10-142-419-222	Sequence 222, App
38	68.5	48.6	290	9 US-10-123-262-222	Sequence 222, App
39	68.5	48.6	290	9 US-10-142-423-222	Sequence 222, App
40	68.5	48.6	290	9 US-10-121-050-222	Sequence 222, App
41	68.5	48.6	290	9 US-10-141-755-222	Sequence 222, App
42	68.5	48.6	290	9 US-10-143-032-222	Sequence 222, App
43	68.5	48.6	290	9 US-10-123-108-222	Sequence 222, App
44	68.5	48.6	290	9 US-10-123-236-222	Sequence 222, App
45	68.5	48.6	290	9 US-10-123-261-222	Sequence 222, App

ALIGNMENTS

RESULT 1
US-10-050-692-5
; Sequence 5, Application US/10050692
; Publication No. US20020182205A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Simmons, David J.
; APPLICANT: Yang, Jinping
; APPLICANT: Redin, William R.
; TITLE OF INVENTION: STIMULATION OF BONE GROWTH WITH THROMBIN
; FILE REFERENCE: 3033.1002-004
; CURRENT APPLICATION NUMBER: US/10/050,692
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 09/909,122
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/219,300
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human prothrombin
US-10-050-692-5

Query Match 100.0%; Score 141; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.5e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTRYKPDGKRGDACEGDSGGPFV 25
|||||
Db 1 AGTRYKPDGKRGDACEGDSGGPFV 25

RESULT 2
US-09-909-348-5
; Sequence 5, Application US/0909348
; Patent No. US20020042373A1

```
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Stierberg, Janet
; APPLICANT: Bergmann, John
; TITLE OF INVENTION: Stimulation Of Cartilage Growth With Agonists
; FILE REFERENCE: 3033.1003-001
; CURRENT APPLICATION NUMBER: US/09/909,348
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/219,800
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide fragment of Thrombin
US-09-909-348-5

Query Match      100.0%; Score 141; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.5e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTRYKPDGKRGDACEGDSGGPFV 25
   |||||
DB 1 AGTRYKPDGKRGDACEGDSGGPFV 25

RESULT 3
US-09-909-122-5
; Sequence 5, Application US/09909122
; Patent No. US20020128202A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Simmons, David J.
; APPLICANT: Yang, Jinping
; APPLICANT: Redin, William R.
; TITLE OF INVENTION: Stimulation Of Bone Growth With Thrombin
; FILE REFERENCE: 3033.1002-001
; CURRENT APPLICATION NUMBER: US/09/909,122
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/219,300
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide fragment of Thrombin
US-09-909-122-5

Query Match      100.0%; Score 141; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.5e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTRYKPDGKRGDACEGDSGGPFV 25
   |||||
DB 1 AGTRYKPDGKRGDACEGDSGGPFV 25

RESULT 4
US-10-050-611-3
; Sequence 3, Application US/10050611
; Publication No. US20020187933A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
```

```
; TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: 3033.1000-008
; CURRENT APPLICATION NUMBER: US/10/050,611
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 09/904,090
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/217,583
; PRIOR FILING DATE: 2000-07-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human fragment of thrombin
US-10-050-611-3

Query Match      85.8%; Score 121; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
   |||||
DB 3 YKPDGKRGDACEGDSGGPFV 23

RESULT 5
US-10-050-611-4
; Sequence 4, Application US/10050611
; Publication No. US20020187933A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
; FILE REFERENCE: 3033.1000-008
; CURRENT APPLICATION NUMBER: US/10/050,611
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 09/904,090
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/217,583
; PRIOR FILING DATE: 2000-07-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: c-terminal amidated fragment of human thrombin
; NAME/KEY: AMIDATION
; LOCATION: 23
; OTHER INFORMATION: valine is amidated as CONH2
US-10-050-611-4

Query Match      85.8%; Score 121; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
   |||||
DB 3 YKPDGKRGDACEGDSGGPFV 23

RESULT 6
US-10-050-688-5
; Sequence 5, Application US/10050688
; Publication No. US20020198154A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
```

APPLICANT: Stiernberg, Janet
TITLE OF INVENTION: STIMULATION OF CARTILAGE GROWTH WITH
TITLE OF INVENTION: AGONISTS OF THE NON-PROTEOLYTICALLY ACTIVATED THROMBIN
TITLE OF INVENTION: RECEPTOR
FILE REFERENCE: 3033.1003-004
CURRENT APPLICATION NUMBER: US/10/050,688
CURRENT FILING DATE: 2002-01-16
PRIOR APPLICATION NUMBER: 09/909,348
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/219,800
PRIOR FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptide fragment of thrombin
US-10-050-688-5

Query Match 85.8%; Score 121; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
Db 3 YKPDGKRGDACEGDSGGPFV 23

RESULT 7

US-10-050-688-6
Sequence 6, Application US/10050688
Publication No. US20020198154A1
GENERAL INFORMATION:
APPLICANT: Carney, Darrell H.
APPLICANT: Crowther, Roger S.
APPLICANT: Stiernberg, Janet
APPLICANT: Bergmann, John
TITLE OF INVENTION: STIMULATION OF CARTILAGE GROWTH WITH
TITLE OF INVENTION: AGONISTS OF THE NON-PROTEOLYTICALLY ACTIVATED THROMBIN
TITLE OF INVENTION: RECEPTOR
FILE REFERENCE: 3033.1003-004
CURRENT APPLICATION NUMBER: US/10/050,688
CURRENT FILING DATE: 2002-01-16
PRIOR APPLICATION NUMBER: 09/909,348
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/219,800
PRIOR FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptide fragment of thrombin
NAME/KEY: AMIDATION
LOCATION: (23)...(23)
OTHER INFORMATION: CONH2
US-10-050-688-6

Query Match 85.8%; Score 121; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
Db 3 YKPDGKRGDACEGDSGGPFV 23

RESULT 8

US-09-777-328-8
Sequence 8, Application US/09777328
Patent No. US20020032314A1
GENERAL INFORMATION:
APPLICANT: CAREY, DARRELL H.
APPLICANT: RAMAKRISHNAN, SHYAM
TITLE OF INVENTION: SYNTHETIC PEPTIDE NEUTROPHIL CELL CHEMOTACTIC AGENTS
FILE REFERENCE: CHBP:002
CURRENT APPLICATION NUMBER: US/09/777,328
CURRENT FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 08/330,594
PRIOR FILING DATE: 1994-10-28
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
US-09-777-328-8

Query Match 85.8%; Score 121; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
Db 3 YKPDGKRGDACEGDSGGPFV 23

RESULT 9

US-09-904-090-3
Sequence 3, Application US/09904090
Patent No. US20020061852A1
GENERAL INFORMATION:
APPLICANT: Carney, Darrell
TITLE OF INVENTION: METHODS OF THERAPY WITH THROMBIN DERIVED
TITLE OF INVENTION: PEPTIDES
FILE REFERENCE: 3033.1000-001
CURRENT APPLICATION NUMBER: US/09/904,090
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: US 60/217,583
PRIOR FILING DATE: 2000-07-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Polypeptide, fragment of thrombin
US-09-904-090-3

Query Match 85.8%; Score 121; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
Db 3 YKPDGKRGDACEGDSGGPFV 23

RESULT 10

US-10-050-692-6
Sequence 6, Application US/10050692
Publication No. US2002018205A1
GENERAL INFORMATION:
APPLICANT: Carney, Darrell H.
APPLICANT: Crowther, Roger S.
APPLICANT: Simmons, David J.

APPLICANT: Yang, Jinning
APPLICANT: Redin, William R.
TITLE OF INVENTION: STIMULATION OF BONE GROWTH WITH THROMBIN
FILE OF INVENTION: PEPTIDE DERIVATIVES
FILE REFERENCE: 3033.1002-004
CURRENT APPLICATION NUMBER: US/10/050,692
CURRENT FILING DATE: 2002-01-16
PRIOR APPLICATION NUMBER: 09/909,122
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: 60/219,300
PRIOR FILING DATE: 2000-07-19
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: c-terminal amidated fragment of human thrombin
NAME/KEY: AMIDATION
LOCATION: (23)...(23)
OTHER INFORMATION: valine is amidated as CONH2
US-10-050-692-6

Query Match 83.7%; Score 118; DB 9; Length 23;
Best Local Similarity 92.0%; Pred. No. 1.1e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 AGTRVYKPDGKRGDACEGDSGGPFV 25
DB 1 AGT--KPDEGKRGDACEGDSGGPFV 23
||| |||||

RESULT 11

US-09-898-837A-45
Sequence 45, Application US/09898837A
Publication No. US20030077697A1
GENERAL INFORMATION:

APPLICANT: Quinn, Kerry E.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Majumder, Kumud
APPLICANT: Vernet, Corine
APPLICANT: Herrmann, John L.
APPLICANT: Burgess, Catherine
APPLICANT: Fernandes, Elma
APPLICANT: Taupier Jr., Raymond
APPLICANT: Rastelli, Luca
APPLICANT: Curagen Corporation
APPLICANT: Gerlach, Valerie L.
APPLICANT: MacDougall, John R.
TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME
FILE REFERENCE: 15966-598 CIP
CURRENT APPLICATION NUMBER: US/09/898,837A
CURRENT FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986
PRIOR FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347
PRIOR FILING DATE: 2000-02-09
PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195
PRIOR FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427
PRIOR FILING DATE: 2000-11-16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 251
TYPE: PRT
ORGANISM: Bos taurus
US-09-898-837A-41

Query Match 80.9%; Score 114; DB 9; Length 251;
Best Local Similarity 95.2%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
DB 191 YKPDGKRGDACEGDSGGPFV 211
||| |||||

PRIOR FILING DATE: 2000-11-16
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 45
LENGTH: 250
TYPE: PRT
ORGANISM: Bos taurus
US-09-898-837A-45

Query Match 80.9%; Score 114; DB 9; Length 250;
Best Local Similarity 95.2%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
DB 190 YKPDGKRGDACEGDSGGPFV 210
||| |||||

RESULT 12

US-09-898-837A-41
Sequence 41, Application US/09898837A
Publication No. US20030077697A1
GENERAL INFORMATION:

APPLICANT: Quinn, Kerry E.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Majumder, Kumud
APPLICANT: Vernet, Corine
APPLICANT: Herrmann, John L.
APPLICANT: Burgess, Catherine
APPLICANT: Fernandes, Elma
APPLICANT: Taupier Jr., Raymond
APPLICANT: Rastelli, Luca
APPLICANT: Curagen Corporation
APPLICANT: Gerlach, Valerie L.
APPLICANT: MacDougall, John R.
TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME
FILE REFERENCE: 15966-598 CIP
CURRENT APPLICATION NUMBER: US/09/898,837A
CURRENT FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986
PRIOR FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347
PRIOR FILING DATE: 2000-02-09
PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195
PRIOR FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427
PRIOR FILING DATE: 2000-11-16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 251
TYPE: PRT
ORGANISM: Bos taurus
US-09-898-837A-41

Query Match 80.9%; Score 114; DB 9; Length 251;
Best Local Similarity 95.2%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:33:02 ; Search time 28.5 Seconds
(without alignments)
84.328 Million cell updates/sec

Title: US-09-909-348-5

Perfect score: 141

Sequence: 1 AGTRYKPDGKRGDACEGDSGGPFV 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	121	85.8	622	1 TBHU	thrombin (EC 3.4.2
2	117	83.0	236	2 C42696	thrombin (EC 3.4.2
3	114	80.9	625	1 TBBO	thrombin (EC 3.4.2
4	108	76.6	234	2 F42696	thrombin (EC 3.4.2
5	103	73.0	235	2 D42696	thrombin (EC 3.4.2
6	103	73.0	235	2 E42696	thrombin (EC 3.4.2
7	100	70.9	236	2 I42696	thrombin (EC 3.4.2
8	99	70.2	239	2 G42696	thrombin (EC 3.4.2
9	92	65.2	617	2 S10511	thrombin (EC 3.4.2
10	92	65.2	618	2 A35827	thrombin (EC 3.4.2
11	79	56.0	235	2 H42696	thrombin (EC 3.4.2
12	71	50.4	461	1 KXHU	protein C (activat
13	70.5	50.0	264	2 S32794	trypsin-like prote
14	67	47.5	225	2 S45356	probable serine pr
15	67	47.5	417	1 S00845	hepsin (EC 3.4.21
16	67	47.5	1524	2 T30337	polyprotein - Afri
17	65	46.1	275	2 S40007	trypsin (EC 3.4.21
18	65	46.1	482	1 EXRT	coagulation factor
19	64.5	45.7	161	2 I48158	coagulation factor
20	64	45.4	191	2 S41115	complement factor
21	64	45.4	246	1 DBHU	complement factor
22	64	45.4	309	2 B49878	coagulation factor
23	64	45.4	456	1 KXBO	protein C (activat
24	64	45.4	1004	2 T30338	oviductin (EC 3.4
25	64	45.4	2616	2 A57096	nudel protein prec
26	63	44.7	161	2 I62744	coagulation factor
27	63	44.7	461	1 JX0210	protein C (activat
28	63	44.7	488	1 EXHU	coagulation factor
29	62	44.0	263	1 I55608	complement factor

30	62	44.0	274	2 JC4171	trypsin (EC 3.4.2
31	62	44.0	461	1 S18994	protein C (activat
32	62	44.0	475	1 EXCH	coagulation factor
33	62	44.0	812	1 PLBO	plasmin (EC 3.4.21
34	61.5	43.6	260	2 A37938	tissue kallikrein
35	61	43.3	285	2 T35195	probable serine pr
36	61	43.3	375	1 A23689	limulus clotting e
37	61	43.3	562	1 UKHUT	t-plasminogen acti
38	60.5	42.9	282	2 I84621	coagulation factor
39	60.5	42.9	392	1 A30100	serine proteinase
40	60.5	42.9	459	2 J00419	coagulation factor
41	60.5	42.9	492	1 EXBO	coagulation factor
42	60.5	42.9	638	1 KOHUP	plasma kallikrein
43	60	42.6	237	1 TRCYL	trypsin (EC 3.4.21
44	60	42.6	250	2 S55493	serine proteinase
45	60	42.6	254	2 S65465	trypsin-like prote

ALIGNMENTS

RESULT 1

TBHU

thrombin (EC 3.4.21.5) precursor [validated] - human

N:Alternate names: coagulation factor II

N:Contains: prothrombin

C:Species: Homo sapiens (man)

C>Date: 30-Nov-1980 #sequence_revision 22-Jul-1994 #text_change 08-Dec-2000

C:Accession: A29351; A00914; B00914; A37549; A37550; I51952

R:Degen, S.J.F.; Davie, E.W.

Biochemistry 26, 6165-6177, 1987

A:Title: Nucleotide sequence of the gene for human prothrombin.

A:Reference number: A29351; MUID:88077877; PMID:2825773

A:Accession: A29351

A:Molecule type: DNA

A:Residues: 1-622 <DE3>

A:Cross-references: GB:M17262; GB:M33691; NID:G558069; PIDN:ARC63054.1; PID:G339641

R:Degen, S.J.F.; Macgillivray, R.T.A.; Davie, E.W.

Biochemistry 22, 2087-2097, 1983

A:Title: Characterization of the complementary deoxyribonucleic acid and gene coding

A:Reference number: A00914; MUID:83231469; PMID:6305407

A:Accession: A00914

A:Molecule type: mRNA

A:Residues: 8-163, 'N', 165-622 <DE2>

A:Cross-references: GB:V00595; GB:J00307; NID:G37128; PIDN:CAA23842.1; PID:G1335344

A:Accession: B00914

A:Molecule type: DNA

A:Residues: 188-311 <DE3>

R:Walz, D.A.; Hewett-Emmett, D.; Seegers, W.H.

Proc. Natl. Acad. Sci. U.S.A. 74, 1969-1972, 1977

A:Reference number: A37549; MUID:77193964; PMID:266717

A:Accession: A37549

A:Molecule type: protein

A:Residues: 44-118, 'N', 120, 'S', 122-163, 'I', 165-175, 'A', 177-182, 'T', 184-193, 'MV', 196-3

R:Butkowski, R.J.; Ellison, J.; Downing, M.R.; Mann, K.G.

J. Biol. Chem. 252, 4942-4957, 1977

A:Title: Primary structure of human prothrombin 2 and alpha-thrombin.

A:Reference number: A37550; MUID:77207112; PMID:873923

A:Accession: A37550

A:Molecule type: protein

A:Residues: 315-334, 'N', 336-348, 'N', 350-368, 'N', 370-397, 'N', 399-413, 'N', 415-484, 'N', 4

R:Rabiet, M.J.; Blashill, A.; Furie, B.; Furie, B.C.

J. Biol. Chem. 261, 13210-13215, 1986

A:Reference number: A37551; MUID:87008532; PMID:3759958

A:Contents: annotation; activation cleavages

R:Macgillivray, R.T.; Irwin, D.M.; Guinto, E.R.; Stone, J.C.

Ann. N. Y. Acad. Sci. 485, 73-79, 1986

A:Title: Recombinant genetic approaches to functional mapping of thrombin.

A:Reference number: I51952; MUID:87182874; PMID:3471151

A:Accession: I51952

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2, 'RI', 5-100 <RES>

A:Cross-references: GB:M33031; NID:q190723; PIDN:AAA60220.1; PID:q190724
 C:Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fibrin
 C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that binds
 ter 314-Arg, are released in natural blood clotting.
 C:Comment: The cleavage after Arg-198, observed in vitro, does not occur in plasma.
 C:Comment: The gamma-carboxyglutamyl residues bind calcium ions, result from the carboxy
 ent interaction with the negatively charged phospholipid membrane surface.
 C:Comment: The prothrombin precursor is synthesized in the liver.
 C:Genetics:
 A:Gene: GDB:F2
 A:Cross-references: GDB:119894; OMIM:176930
 A:Map position: 11p11-11q12
 A:Introns: 27/1; 80/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2; 433/2; 491/2; 552
 C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
 C:Keywords: acute phase; blood coagulation; calcium binding; carboxyglutamic acid; dupli
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-43/Domain: propeptide #status predicted <PRO>
 F:28-87/Domain: Gla domain homology <GLA>
 F:44-622/Product: prothrombin #status experimental <MAT>
 F:108-186/Domain: activation peptide #status experimental <APT>
 F:213-291/Domain: kringle homology <KR1>
 F:328-363/Product: thrombin light chain #status experimental <LCH>
 F:364-622/Product: thrombin heavy chain #status experimental <HCH>
 F:364-613/Domain: trypsin homology <TRY>
 F:49-50-57-59-62-63-68-69-72-75/Modified site: gamma-carboxyglutamic acid (Glu) #status
 F:60-65-90-103-108-186-129-169-157-181-213-291-234-274-262-286/Disulfide bonds: #status
 F:121-143/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:336-482-536-550-564-594/Disulfide bonds: #status predicted
 F:391-407/Disulfide bonds: #status experimental
 F:406-462/Active site: His, Asp #status predicted
 F:416/Binding site: carboxylate (Asn) (covalent) #status experimental
 F:568/Active site: Ser #status experimental

Query Match 85.8%; Score 121; DB 1; Length 622;
 Best Local Similarity 100.0%; Pred. No. 5.9e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGSGGPFV 25
 |||
 Db 553 YKPDGKRGDACEGSGGPFV 573

RESULT 2
 C42696
 thrombin (EC 3.4.21.5) B chain - rabbit (fragment)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
 C:Accession: C42696
 R:Banfield, D.K.; MacGillivray, R.T.A.
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
 A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq
 A:Reference number: A42696; MUID:94212913; PMID:1557383
 A:Accession: C42696
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA
 A:Residues: 1-236 <BAN>
 A:Cross-references: GB:M81396
 C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
 C:Keywords: hydrolase; serine proteinase
 F:1-227/Domain: trypsin homology (fragment) <TRY>

Query Match 83.0%; Score 117; DB 2; Length 236;
 Best Local Similarity 95.2%; Pred. No. 8.3e-09;
 Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGSGGPFV 25
 |||
 Db 167 YKPDGKRGDACEGSGGPFV 187

RESULT 3

TBEO
 thrombin (EC 3.4.21.5) precursor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 24-Apr-1994 #sequence_revision 14-Jul-1994 #text_change 18-Jun-1999
 C:Accession: S02537; A00915; A37552; I46045; S67518
 R:Irwin, D.M.; Robertson, K.A.; MacGillivray, R.T.A.
 J. Mol. Biol. 200, 31-45, 1988
 A:Title: Structure and evolution of the bovine prothrombin gene.
 A:Reference number: S02537; MUID:88245190; PMID:3379642
 A:Accession: S02537
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-625 <IRW>
 R:MacGillivray, R.T.A.; Davie, E.W.
 Biochemistry 23, 1626-1634, 1984
 A:Title: Characterization of bovine prothrombin mRNA and its translation product.
 A:Reference number: A00915; MUID:84203525; PMID:6326805
 A:Accession: A00915
 A:Molecule type: mRNA
 A:Residues: 1-230, 'H', 232-625 <MAC>
 A:Note: 600-Asn was also found
 R:Magnusson, S.; Sottrup-Jensen, L.; Petersen, T.E.; Claeyss, H.
 In Boerhaave Symposium on Prothrombin and Related Coagulation Factors, Hemker, H.C.,
 A:Reference number: A37552
 A:Accession: A37552
 A:Molecule type: protein
 A:Residues: 44-287, 'N', 289-352, 'E', 354, 'Q', 356-548, 'ND', 551-599, 'N', 601-625 <MAG>
 A:Note: the evidence for 231-Ser is strong
 A:Note: disulfide bonds and carbohydrate binding sites were determined
 R:Park, C.H.; Tullinsky, A.
 Biochemistry 25, 3977-3982, 1986
 A:Title: Three-dimensional structure of the kringle sequence: structure of prothrombi
 A:Reference number: A37553; MUID:86296631; PMID:3741841
 A:Contents: annotation; residues 44-317, X-ray crystallography, 2.8 angstroms
 R:Irwin, D.M.; Ahern, K.G.; Pearson, G.D.; MacGillivray, R.T.A.
 Biochemistry 24, 6854-6861, 1985
 A:Title: Characterization of the bovine prothrombin gene.
 A:Reference number: A37554; MUID:86077733; PMID:3000440
 A:Contents: annotation; gene structure
 R:MacGillivray, R.T.; Degen, S.J.; Chandra, T.; Woo, S.L.; Davie, E.W.
 Proc. Natl. Acad. Sci. U.S.A. 77, 5153-5157, 1980
 A:Title: Cloning and analysis of a cDNA coding for bovine prothrombin.
 A:Reference number: I46045; MUID:81054926; PMID:6254059
 A:Accession: I46045
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 466-599, 'N', 601-625 <MA2>
 A:Cross-references: EMBL:V00135; NID:q772; PIDN:CAA23451.1; PID:9808945
 R:Pejler, G.; Karlstrom, A.R.; Berg, L.
 Eur. J. Biochem. 227, 102-107, 1995
 A:Title: Identification of the proteolytic thrombin fragments formed after cleavage w
 A:Reference number: S67518; MUID:95154277; PMID:7851376
 A:Accession: S67518
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 318-325; 333-338, 'X', 340; 367-374; 481-484, 'X', 486-488; 515-522 <PEJ>
 C:Comment: Thrombin, which cleaves bonds after Arg and Lys, converts fibrinogen to fi
 C:Comment: Prothrombin is activated on the surface of a phospholipid membrane that bi
 tivation peptide and cleaves the remaining part into light and heavy chains. The acti
 C:Comment: Thrombin can cleave the amino-terminal activation peptide 1 from prothromb
 C:Comment: The gamma-carboxyglutamyl residues bind calcium ions, result from the carb
 ent interaction with the negatively charged phospholipid membrane surface.
 C:Comment: The prothrombin precursor is synthesized in the liver.
 C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
 C:Keywords: blood coagulation; calcium binding; carboxyglutamic acid; duplication; gl
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-43/Domain: propeptide #status predicted <PRO>
 F:28-88/Domain: Gla domain homology <GLA>
 F:44-625/Product: prothrombin #status experimental <MPT>
 F:44-199/Domain: activation peptide 1 #status experimental <PR1>
 F:109-187/Domain: kringle homology <KR1>
 F:200-317/Domain: activation peptide 2 #status experimental <PR2>

RESULT 6
E42696
thrombin (EC 3.4.21.5) B chain - tokay (fragment)
C:Species: Gekko gekko (tokay)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
C:Accession: E42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: E42696
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-235 <BAN>
A:Cross-references: GB:M81392
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 73.0%; Score 103; DB 2; Length 235;
Best Local Similarity 81.0%; Pred. No. 7.1e-07;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
| ::| ||||| ||||| |||||
DB 166 YSPEDSKRGDACEGDSGGPFV 186

RESULT 7
I42696
thrombin (EC 3.4.21.5) B chain - Pacific hagfish (fragment)
C:Species: Eptatretus stouti (Pacific hagfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
C:Accession: I42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: I42696
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-236 <BAN>
A:Cross-references: GB:M81393
A:Note: nucleotide translation not given
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 70.9%; Score 100; DB 2; Length 236;
Best Local Similarity 81.0%; Pred. No. 1.8e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
| ::| ||||| ||||| |||||
DB 166 YSPEDSKRGDACEGDSGGPFV 186

RESULT 8
G42696
thrombin (EC 3.4.21.5) B chain - rainbow trout (fragment)
C:Species: Oncorhynchus mykiss (rainbow trout)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: G42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: G42696
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-239 <BAN>

C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
 C:Keywords: hydrolase; serine proteinase
 F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 70.2%; Score 99; DB 2; Length 239;
 Best Local Similarity 81.0%; Pred. No. 2.6e-06;
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
 :||| | ||||| |||||
 Db 166 FRPEQKRGDACEGDSGGPFV 186

RESULT 9

S10511
 thrombin (EC 3.4.21.5) precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 03-May-2002
 C:Accession: S10511; A60576; B42696
 R:Dibanich, M.; Monard, D.
 Nucleic Acids Res. 18, 4251, 1990
 A:Title: cDNA sequence of rat prothrombin.
 A:Reference number: S10511; MUID:90332426; PMID:2377469
 A:Accession: S10511
 A:Molecule type: mRNA
 A:Residues: 1-617 <DIB>
 A:Cross-references: EMBL:X52835; NID:956969; PIDN:CAA37017.1; PID:956970
 R:Henriksson, K.P.; Jasin, E.E.; Greenwood, J.A.; Dickerman, H.W.
 Endocrinology 126, 167-175, 1990
 A:Title: Prothrombin levels are increased in the estrogen-treated immature rat uterus.
 A:Reference number: A60576; MUID:90091942; PMID:2293980
 A:Accession: A60576
 A:Molecule type: protein
 A:Residues: 44-58 <HEN>
 A:Note: the authors purified the proenzyme from the estrogen-stimulated maturing rat uterus
 R:Banfield, D.K.; MacGillivray, R.T.A.
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
 A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq

A:Reference number: A42696; MUID:92212913; PMID:1557383
 A:Accession: B42696
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 383-617, 'E' <BAN>
 A:Cross-references: GB:M81397
 C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
 C:Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; hyd

F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-43/Domain: propeptide #status predicted <PRO>
 F:44-617/Product: prothrombin #status experimental <PMAT>
 F:109-187/Domain: kringle homology <KR1>
 F:215-292/Domain: kringle homology <KR2>
 F:360-603/Domain: trypsin homology <TRY>
 F:501-51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxylglutamic acid (Glu) #status
 F:61-66,91-104,109-187,130-170,158-182,215-292,236-276,264-287,332-478,387-403,532-546,5
 F:402,458,564/Active site: His, Asp, Ser #status predicted

Query Match 65.2%; Score 92; DB 2; Length 617;
 Best Local Similarity 76.2%; Pred. No. 5.9e-05;
 Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
 :||| | ||||| |||||
 Db 549 FRVNDTKRGDACEGDSGGPFV 569

RESULT 10

A35827
 thrombin (EC 3.4.21.5) precursor - mouse
 C:Species: Mus musculus (House mouse)
 C:Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 03-May-2002
 C:Accession: A35827; A42696; S12081
 R:Degen, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pai, J.A.

DNA Cell Biol. 9, 487-498, 1990
 A:Title: Characterization of the cDNA coding for mouse prothrombin and localization o
 A:Reference number: A35827; MUID:91025551; PMID:2222810
 A:Accession: A35827
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-618 <DEG>
 A:Cross-references: GB:X52308; NID:953813; PIDN:CAA36548.1; PID:953814
 A:Experimental source: strain C57BL/6
 A:Note: the data were obtained from females resulting from the cross of M. domesticus
 R:Banfield, D.K.; MacGillivray, R.T.A.
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
 A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
 A:Reference number: A42696; MUID:92212913; PMID:1557383
 A:Accession: A42696
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 384-618, 'B' <BAN>
 A:Cross-references: GB:M81394
 C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
 C:Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; h

F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-43/Domain: propeptide #status predicted <PRO>
 F:28-88/Domain: Gla domain homology <GLA>
 F:44-618/Product: prothrombin B #status predicted <MAT>
 F:109-187/Domain: kringle homology <KR1>
 F:215-293/Domain: kringle homology <KR2>
 F:361-610/Domain: trypsin homology <TRY>
 F:501-51,58,60,63,64,69,70,73,76/Modified site: gamma-carboxylglutamic acid (Glu) #stat
 F:61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-54
 F:403,459,565/Active site: His, Asp, Ser #status predicted

Query Match 65.2%; Score 92; DB 2; Length 618;
 Best Local Similarity 76.2%; Pred. No. 5.9e-05;
 Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
 :||| | ||||| |||||
 Db 550 FKVNDTKRGDACEGDSGGPFV 570

RESULT 11

H42696
 thrombin (EC 3.4.21.5) B chain - white sturgeon (fragment)
 C:Species: Acipenser transmontanus (White sturgeon)
 C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
 A:Accession: H42696
 R:Banfield, D.K.; MacGillivray, R.T.A.
 Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
 A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and
 A:Reference number: A42696; MUID:92212913; PMID:1557383
 A:Accession: H42696
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 1-235 <BAN>
 A:Cross-references: GB:M81399
 C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
 C:Keywords: hydrolase; serine proteinase
 F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 56.0%; Score 79; DB 2; Length 235;
 Best Local Similarity 61.9%; Pred. No. 0.0015;
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 YKPDEGKRGDACEGDSGGPFV 25
 :||| | ||||| |||||
 Db 166 FSPEDSISGACEDSGSGPFV 186

RESULT 12

KXHU
 protein C (activated) (EC 3.4.21.69) precursor - human
 N:Alternate names: autoprothrombin IIA; plasma protein C

C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
C;Accession: A22331; A2426; A21781; A23789; A00927
R;Foster, D.C.; Yoshitake, S.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 4673-4677, 1985
A;Title: The nucleotide sequence of the gene for human protein C.
A;Reference number: A22331; MUID:85270390; PMID:2991887
A;Accession: A22331
A;Molecule type: DNA
A;Residues: 1-461 <FOS1>
A;Cross-references: GB:M1128; NID:g190333; PIDN:AAA60166.1; PID:g190334
R;Plutsky, J.; Hoskins, J.A.; Long, G.L.; Crabtree, G.R.
Proc. Natl. Acad. Sci. U.S.A. 83, 546-550, 1986
A;Title: Evolution and organization of the human protein C gene.
A;Reference number: A25426; MUID:86120978; PMID:3511471
A;Accession: A25426
A;Molecule type: DNA
A;Residues: 1-445, 'L', 446-461 <PLD>
A;Cross-references: GB:M12712; NID:g190330; PIDN:AAA60165.1; PID:g190332
R;Foster, D.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 4766-4770, 1984
A;Title: Characterization of a cDNA coding for human protein C.
A;Reference number: A21781; MUID:84272714; PMID:6589623
A;Accession: A21781
A;Molecule type: mRNA
A;Residues: 'Q', 107-461 <FOS2>
A;Cross-references: GB:R02059; NID:g190322; PIDN:AAA60164.1; PID:g190323
R;Beckmann, R.J.; Schmidt, R.J.; Santerre, R.F.; Plutsky, J.; Crabtree, G.R.; Long, G.L.
Nucleic Acids Res. 13, 5233-5247, 1985
A;Title: The structure and evolution of a 461 amino acid human protein C precursor and its processing.
A;Reference number: A23789; MUID:85269639; PMID:2991859
A;Accession: A23789
A;Molecule type: mRNA
A;Residues: 1-461 <BEC>
A;Cross-references: GB:R02750; NID:g35689; PIDN:CRA26528.1; PID:g763120
R;Miletich, J.P.; Broze Jr., G.J.
J. Biol. Chem. 265, 11397-11404, 1990
A;Title: Beta protein C is not glycosylated at asparagine 329. The rate of translation m
A;Reference number: A44605; MUID:90293094; PMID:1694179
A;Contents: annotation; carbohydrate binding sites; activation peptide
A;Note: the alpha form of protein C is glycosylated at Asn-329, and the beta form is not
R;Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A;Title: O-linked fucose is present in the first epidermal growth factor domain of facto
A;Reference number: A44606; MUID:92184750; PMID:1544894
A;Contents: annotation; beta-hydroxyaspartic acid
C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that in
activation of factor Va is strongly enhanced by complexing with protein S. Protein C also in
C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is d
bin, which cleaves a dodecapeptide from the amino end of the heavy chain; this reaction,
C;Genetics:
A;Gene: GDB:PROC
A;Cross-references: GDB:120317; OMIM:176860
A;Map position: 2q13-2q21
A;Introns: 24/1; 79/3; 88/1; 134/1; 179/1; 226/3; 266/1
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding
F:1-32/Domain: signal sequence #status predicted <SIG>
F:27-86/Domain: Gla domain homology <GLA>
F:33-42/Domain: propeptide #status predicted <PRO>
F:43-197/Product: protein C light chain #status predicted <LCH>
F:92-131/Domain: EGF homology <EG2>
F:140-175/Domain: EGF homology <EG1>
F:200-461/Product: protein C heavy chain #status predicted <HCH>
F:200-211/Product: activation peptide #status experimental <APT>
F:212-445/Domain: trypsin homology <TRY>
F:48,49,56,58,61,62,67,68,71/Modified site: gamma-carboxyglutamic acid (Glu) #status exp
F:59-64,92,105,107-120,121-131,140-151,147-160,162-175,183-319,238-254,373-387,398-426/H
F:106-111/Disulfide bonds: #status predicted
F:110/Binding site: carboxylate (Thr) (covalent) #status absent
F:113/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
F:139,290,355/Binding site: carboxylate (Asn) (covalent) #status experimental
F:211-212/Cleavage site: Arg-Leu (thrombin) #status experimental

```

F:253,299,402/Active site: His, Asp, Ser #status predicted
F:371/Binding site: carbohydrate (Asn) (covalent) (partial) #status atypical

Query Match          50.4%; Score 71; DB 1; Length 461;
Best Local Similarity 81.2%; Pred. No. 0.035;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 10 GKRGDACEGDSGGPFV 25
      ||| ||||| ||||| |||
Db 392 GDRQDACEGDSGGPMV 407

RESULT 13
S32794
trypsin-like proteinase (EC 3.4.21.-) - silkworm
C:Species: Bombyx mori (silkworm)
C:Date: 07-Dec-1994 #sequence_revision 01-Sep-1995 #text_change 11-May-2000
C:Accession: S32794
R:Ikeda, M.; Yaginuma, T.; Kobayashi, M.; Yamashita, O.
Comp. Biochem. Physiol. B 99, 405-411, 1991
A:Title: cDNA cloning, sequencing and temporal expression of the protease responsible
A:Reference number: S32794; MUID:92111263; PMID:1764920
A:Accession: S32794
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-264 <IKEX>
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:28-248/Domain: trypsin homology <TRY>

Query Match          50.0%; Score 70.5; DB 2; Length 264;
Best Local Similarity 64.0%; Pred. No. 0.024;
Matches 16; Conservative 2; Mismatches 2; Indels 5; Gaps 2;

QY 1 AGTRYKPDGKRGDACEGDSGGPFV 25
      ||| | : || ||||| |||
Db 195 AGT---PEGGK--DACQDGGGLV 214

RESULT 14
S45356
Probable serine proteinase (EC 3.4.21.-) ACO - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 14-Apr-1998 #sequence_revision 08-May-1998 #text_change 19-May-2000
C:Accession: S45356
R:Dihanich, M.; Spieles, M.
Biochim. Biophys. Acta 1218, 225-228, 1994
A:Title: A novel serine proteinase-like sequence from human brain.
A:Reference number: S45356; MUID:94289486; PMID:8018728
A:Accession: S45356
A:Molecule type: mRNA
A:Residues: 1-225 <DIH>
A:Cross-references: EMBL:X75363; NID:g407137; PIDN:CAA53145.1; PID:g940540
A:Experimental source: Alzheimer's disease patient brain cortex
C:Genetics:
A:Gene: ACO
A:Introns: 175/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase

Query Match          47.5%; Score 67; DB 2; Length 225;
Best Local Similarity 64.7%; Pred. No. 0.063;
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 EGKRGDACEGDSGGPFV 25
      || : ||||| |||
Db 167 EGKAECEGDSGGPLV 183

RESULT 15
S00845
hepsin (EC 3.4.21.-) - human
C:Species: Homo sapiens (man)

```

Search completed: May 12, 2003, 15:39:44
Job time : 28.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:32:07 ; Search time 14.5 seconds
(without alignments)
71.511 Million cell updates/sec

Title: US-09-909-348-5

Perfect score: 141

Sequence: 1 ACTRYKPDGKRGDAGEGDSGGPFV 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	85.8	622	1 THRB_HUMAN	P00734 homo sapien
2	114	80.9	625	1 THRB_BOVIN	P00735 bos taurus
3	92	65.2	617	1 THRB_RAT	P18292 rattus norv
4	92	65.2	618	1 THRB_MOUSE	P19221 mus musculus
5	71	50.4	161	1 PRTC_MACMO	Q28506 macaca mula
6	71	50.4	461	1 PRTC_HUMAN	P04070 homo sapien
7	70.5	50.0	264	1 VDP_BOMMO	Q07943 bombyx mori
8	70	49.6	281	1 TRYZ_DROER	P54630 drosophila
9	68.5	48.6	290	1 MPN_HUMAN	Q9bqr3 homo sapien
10	68	48.2	458	1 PRTC_RABIT	Q28661 oryctolagus
11	67	47.5	256	1 KLKE_HUMAN	Q9h245 homo sapien
12	67	47.5	416	1 HEPS_MOUSE	O35453 mus musculus
13	67	47.5	417	1 HEPS_HUMAN	P05981 homo sapien
14	65.5	46.5	277	1 KLKD_HUMAN	Q9ukr3 homo sapien
15	65	46.1	157	1 PRTC_CANFA	Q28278 canis famil
16	65	46.1	157	1 PRTC_CAPHI	Q28315 capra hircu
17	65	46.1	157	1 PRTC_FELCA	Q28412 felis silve
18	65	46.1	157	1 PRTC_HORSE	Q28380 equus cabal
19	65	46.1	275	1 TRY3_ANOGA	P35091 anopheles g
20	65	46.1	455	1 TMS5_MOUSE	Q9e0r4 mus musculus
21	65	46.1	457	1 TMS5_HUMAN	Q9h3s3 homo sapien
22	65	46.1	459	1 PRTC_PIG	Q9qlp2 sus scrofa
23	64.5	45.7	422	1 DES1_HUMAN	Q9ul52 homo sapien
24	64	45.4	233	1 CFAD_HUMAN	P00746 homo sapien
25	64	45.4	259	1 CFAD_PIG	P17779 sus scrofa
26	64	45.4	456	1 PRTC_BOVIN	P00745 bos taurus
27	64	45.4	566	1 TPA_BOVIN	Q28198 bos taurus
28	64	45.4	875	1 NETR_HUMAN	P56730 homo sapien
29	64	45.4	2616	1 NDLD_DROME	P98159 drosophila
30	63	44.7	256	1 TRYE_DROER	P54627 drosophila
31	63	44.7	461	1 PRTC_MOUSE	P33587 mus musculus
32	63	44.7	488	1 FA10_HUMAN	P00742 homo sapien
33	62	44.0	251	1 KLKE_HUMAN	Q9p0g3 homo sapien

34	62	44.0	263	1 CFAD_RAT	P32038 rattus norv
35	62	44.0	274	1 MCT6_RAT	P50343 rattus norv
36	62	44.0	461	1 PRTC_RAT	P31394 rattus norv
37	62	44.0	475	1 FA10_CHICK	P25155 gallus gall
38	62	44.0	653	1 HGFA_MOUSE	Q9r098 mus musculus
39	62	44.0	812	1 PLMN_BOVIN	P06868 bos taurus
40	61.5	43.6	260	1 ESTA_CANFA	P09582 canis famil
41	61.5	43.6	324	1 TEST_MOUSE	Q9j1b7 mus musculus
42	61.5	43.6	490	1 FA10_RABIT	O19045 oryctolagus
43	61.5	43.6	603	1 CFAD_MOUSE	Q61129 mus musculus
44	61	43.3	255	1 TRY4_LUCCU	P35044 lucilia cup
45	61	43.3	256	1 TRYE_DROME	P35005 drosophila

ALIGNMENTS

RESULT 1
THRB_HUMAN
ID THRB_HUMAN STANDARD; PRT; 622 AA.
AC P00734;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II).
GN P2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88077877; PubMed=2825773;
RA Degen S.J.F., Davie E.W.;
RT "Nucleotide sequence of the gene for human prothrombin.";
RL Biochemistry 26:6165-6177(1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT MET-165.
RA Rieder M.J., Armet T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
Ozuna M., Poel C.L., Toth E.J., Yi O., Nickerson D.A.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 8-622 FROM N.A.
RX MEDLINE=83231459; PubMed=6305407;
RA Degen S.J.F., McGillivray R.T.A., Davie E.W.;
RT "Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin.";
RL Biochemistry 22:2087-2097(1983).
RN [4]
RP SEQUENCE OF 44-314.
RX MEDLINE=77193964; PubMed=266717;
RA Walz D.A., Hewett-Emmett D., Seegers W.H.;
RT "Amino acid sequence of human prothrombin fragments 1 and 2.";
RL Proc. Natl. Acad. Sci. U.S.A. 74:1969-1972(1977).
RN [5]
RP SEQUENCE OF 315-622.
RX MEDLINE=77207112; PubMed=873923;
RA Butkowski R.J., Ellison J., Downing M.R., Mann K.G.;
RT "Primary structure of human prothrombin 2 and alpha-thrombin.";
RL J. Biol. Chem. 252:4942-4957(1977).
RN [6]
RP PROCESSING.
RX MEDLINE=87008532; PubMed=3759958;
RA Rabiet M.J., Blashill A., Furie B., Furie B.C.;
RT "Prothrombin fragment 1 X 2 X 3, a major product of prothrombin activation in human plasma.";
RL J. Biol. Chem. 261:13210-13215(1986).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=90059942; PubMed=2583108;
RA Bode W., Mayr I., Baumann U., Huber R., Stone S.R., Hofsteenge J.;
RT "The refined 1.9 A crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of

the Tyr-Pro-Trp insertion segment.";
EMBO J. 8:3467-3475(1989).
[8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE-90327074; PubMed-2374926;
RA Rydel T.J., Ravichandran K.G., Tulinsky A., Bode W., Huber R.,
RA Roitsch C., Penton J.W. II;
RT "The structure of a complex of recombinant hirudin and human alpha-
RT thrombin.";
RL Science 249:277-280(1990).
[9]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE-94350942; PubMed-8071320;
RA Rydel T.J., Yin M., Padmanabhan K.P., Blankenship D.T., Cardin A.D.,
RA Correa P.E., Fenton J.W. II, Tulinsky A.;
RT "Crystallographic structure of human gamma-thrombin.";
RL J. Biol. Chem. 269:22000-22006(1994).
[10]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE-97357286; PubMed-9214615;
RA van de Locht A., Bode W., Huber R., le Bonniec B.F., Stone S.R.,
RA Esmen C.T., Stubbs M.T.;
RT "The thrombin E192Q-Bpti complex reveals gross structural
RT rearrangements: Implications for the interaction with antithrombin
RT and thrombomodulin.";
RL EMBO J. 16:2977-2984(1997).
[11]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.
RX MEDLINE-99162521; PubMed-10051558;
RA Guinto E.R., Caccia S., Rose T., Fuetterer K., Waksman G., di Cera E.;
RT "Unexpected crucial role of residue 225 in serine proteases.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857(1999).
[12]
RP VARIANT BARCELONA.
RX MEDLINE-87033739; PubMed-3771562;
RA Rabiet M.-J., Furie B.C., Furie B.;
RT "Molecular defect of prothrombin Barcelona. Substitution of cysteine
RT for arginine at residue 273.";
RL J. Biol. Chem. 261:15045-15048(1986).
[13]
RP VARIANT FRANKFURT.
RX MEDLINE-95313001; PubMed-7792730;
RA Degen S.J.F., McDowell S.A., Sparks L.M., Scharer I.;
RT "Prothrombin Frankfurt: a dysfunctional prothrombin characterized by
RT substitution of Glu-466 by Ala.";
RL Thromb. Haemost. 73:203-206(1995).
[14]
RP VARIANTS HIMI-1 AND HIMI-2.
RX MEDLINE-93043342; PubMed-1421398;
RA Morishita E., Saito M., Kumabashiri I., Asakura H., Matsuda T.,
RA Yanaguchi K.;
RT "Prothrombin Himi: a compound heterozygote for two dysfunctional
RT prothrombin molecules (Met-337-->Thr and Arg-388-->His).";
RL Blood 80:2275-2280(1992).
[15]
RP VARIANT PADUA-1.
RX MEDLINE-95169898; PubMed-7865694;
RA James H.L., Kim D.J., Zheng D.-Q., Girolami A.;
RT "Substitution at a factor Xa cleavage site.";
RL Blood Coagul. Fibrinolysis 5:841-844(1994).
[16]
RP VARIANT QUICK-1.
RX MEDLINE-89207504; PubMed-3242619;
RA Henriksen R.A., Mann K.G.;
RT "Identification of the primary structural defect in the dysthrombin
RT thrombin Quick 1: substitution of cysteine for arginine-382.";
RL Biochemistry 27:9160-9165(1988).
[17]
RP VARIANT QUICK-2.
RX MEDLINE-89247398; PubMed-2719846;
RA Henriksen R.A., Mann K.G.;
RT "Substitution of valine for glycine-558 in the congenital dysthrombin

thrombin Quick II alters primary substrate specificity.";
RL Biochemistry 28:2078-2082(1989).
[18]
RP VARIANT SALAKTA.
RX MEDLINE-92378975; PubMed-1354985;
RA Miyata T., Aruga R., Uneyama H., Bezeaud A., Guillin M.-C.,
RA Iwanaga S.;
RT "Prothrombin Salakta: substitution of glutamic acid-466 by alanine
RT reduces the fibrinogen clotting activity and the esterase activity.";
RL Biochemistry 31:7457-7462(1992).
[19]
RP VARIANT TOKUSHIMA.
RX MEDLINE-87185407; PubMed-3567158;
RA Miyata T., Morita T., Inomoto T., Kawauchi S., Shirakami A.,
RA Iwanaga S.;
RT "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan
RT that impairs the fibrinogen clotting activity of derived thrombin
RT Tokushima.";
RL Biochemistry 26:1117-1122(1987).
[20]
RP VARIANT TOKUSHIMA.
RX MEDLINE-87101511; PubMed-3801671;
RA Inomoto T., Shirakami A., Kawauchi S., Shigeakiyo T., Saito S.,
RA Miyoshi K., Morita T., Iwanaga S.;
RT "Prothrombin Tokushima: characterization of dysfunctional thrombin
RT derived from a variant of human prothrombin.";
RL Blood 69:565-569(1987).
[21]
RP VARIANT TOKUSHIMA.
RX MEDLINE-92256895; PubMed-1349838;
RA Iwahana H., Yoshimoto K., Shigeakiyo T., Shirakami A., Saito S.,
RA Itakura M.;
RT "Detection of a single base substitution of the gene for prothrombin
RT Tokushima. The application of PCR-SSCP for the genetic and molecular
RT analysis of dysprothrombinemia.";
RL Int. J. Hematol. 55:93-100(1992).
[22]
RP VARIANT TYPE-3.
RX MEDLINE-83204687; PubMed-6405779;
RA Board P.G., Shaw D.C.;
RT "Determination of the amino acid substitution in human prothrombin
RT type 3 (157 Glu leads to Lys) and the localization of a third
RT thrombin cleavage site.";
RL Br. J. Haematol. 54:245-254(1983).
CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC fibrinogen to fibrin and releases fibrinopeptide A and B.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
CC -1- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOomal
CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
CC OF PROTHROMBIN TO THROMBIN.
CC -1- DISEASE: DEFECTS IN F2 ARE THE CAUSE OF VARIOUS FORMS OF
CC DYSPROTHROMBINEMIA.
CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC THROMBIN.
CC -1- MISCELLANEOUS: IT IS NOT KNOWN WHETHER 1 OR 2 SMALLER ACTIVATION
CC PEPTIDES, WITH ADDITIONAL CLEAVAGE AFTER 314-ARG, ARE RELEASED IN
CC NATURAL BLOOD CLOTTING.
CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC BY FACTOR XA.
CC -1- MISCELLANEOUS: THE CLEAVAGE AFTER R-198, OBSERVED IN VITRO, DOES

Query Match 85.88; Score 121; DB 1; Length 622;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGSGGPFV 25
 |||||
 DB 553 YKPDGKRGDACEGSGGPFV 573

RESULT 2

THRE_BOVIN STANDARD; PRT; 625 AA.
 AC P00735;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Prothrombin precursor (EC 3.4.21.5).
 GN F2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88245190; PubMed-3379642;
 RA Irwin D.M., Robertson K.A., Macgillivray R.T.A.;
 RT "Structure and evolution of the bovine prothrombin gene.";
 RL J. Mol. Biol. 200;31-45(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-84203525; PubMed-6326805;
 RA Macgillivray R.T.A., Davie E.W.;
 RT "Characterization of bovine prothrombin mRNA and its translation product.";
 RL Biochemistry 23:1626-1634(1984).
 RN [3]
 RP SEQUENCE OF 44-625, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
 RA Magnusson S., Sottrup-Jensen L., Petersen T.E., Claeys H.;
 RL (in) Hemker H.C., Velkamp J.J. (eds.);
 RL Boerhaave symposium on prothrombin and related coagulation factors,
 PP.25-46, Leiden University Press, Leiden (1975).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RX MEDLINE-86296631; PubMed-3741841;
 RA Park C.H., Tullinsky A.;
 RT "Three-dimensional structure of the kringle sequence: structure of prothrombin fragment 1.";
 RL Biochemistry 25:3977-3982(1986).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RX MEDLINE-91311686; PubMed-1856869;
 RA Seshadri T.P., Tullinsky A., Skrzypczak-Jankun E., Park C.H.;
 RT "Structure of bovine prothrombin fragment 1 refined at 2.25-A resolution.";
 RL J. Mol. Biol. 220:481-494(1991).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RX MEDLINE-92190185; PubMed-1547238;
 RA Soriano-Garcia M., Padmanabhan K., de Vos A.M., Tullinsky A.;
 RT "The Ca2+ ion and membrane binding structure of the Gla domain of Ca-prothrombin fragment 1.";
 RL Biochemistry 31:2554-2566(1992).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE-92218459; PubMed-1560020;
 RA Martin P.D., Robertson W., Turk D., Huber R., Edwards B.F.P.;
 RT "The structure of residues 7-16 of the A alpha-chain of human fibrinogen bound to bovine thrombin at 2.3-A resolution.";
 RL J. Biol. Chem. 267:7911-7920(1992).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).

RX MEDLINE-92389319; PubMed-1518046;
 RA Brandstetter H., Turk D., Hoeffken H.W., Grosse D., Stuerzebecher J.,
 RA Martin P.D., Edwards B.F.P., Bode W.;
 RT "Refined 2.3 A X-ray crystal structure of bovine thrombin complexes formed with the benzamidine and arginine-based thrombin inhibitors NAPAP, 4-TAPAP and MQPA. A starting point for improving antithrombotics.";
 RT J. Mol. Biol. 226:1085-1089(1992).
 RL [9]
 RN X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.
 RX MEDLINE-97102783; PubMed-8947023;
 RA van de Locht A., Stubbs M.T., Bode W., Friedrich T., Bollschweiler C.,
 RA Hoeffken W., Huber R.;
 RT "The ornithodorin-thrombin crystal structure, a key to the TAP enigma?";
 RT EMBO J. 15:6011-6017(1996).
 RL [10]
 RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH TRIABIN.
 RX MEDLINE-98004486; PubMed-9342325;
 RA Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,
 RA Huber R., Bode W.;
 RT "Structure of the thrombin complex with triabin, a lipocalin-like exosite-binding inhibitor derived from a triatomine bug.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11845-11850(1997).
 RN [11]
 RP GENE STRUCTURE.
 RX MEDLINE-86077733; PubMed-3000440;
 RA Irwin D.M., Ahern K.G., Pearson G.D., Macgillivray R.T.A.;
 RT "Characterization of the bovine prothrombin gene.";
 RL Biochemistry 24:6854-6861(1985).
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII, AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-[1-Gly; activates fibrinogen to fibrin and releases fibrinopeptide A and B.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
 CC -1- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS, RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION OF PROTHROMBIN TO THROMBIN.
 CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN & FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT & HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF THROMBIN.
 CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION BY FACTOR XA.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC -1- DATABASE: NAME-Prozyme technical fact sheet;
 CC WWW-http://www.prozyme.com/technical/thrombindata.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V00135; CAA23451.1;
 CC EMBL; J00041; AAA30781.1;
 CC PIR; A00915; TBBO.
 CC PIR; S02537; S02537.
 CC PDB; 1BBR; 31-JAN-94.
 CC PDB; 1ETR; 31-JAN-94.
 CC PDB; 1ETS; 31-JAN-94.

DR PDB: 1ETT; 31-JAN-94.
 DR PDB: 1HRT; 31-JAN-94.
 DR PDB: 2PFI; 31-JAN-94.
 DR PDB: 2PF2; 31-JAN-94.
 DR PDB: 2SPT; 31-MAY-94.
 DR PDB: 1MKW; 07-JUL-97.
 DR PDB: 1MKX; 07-JUL-97.
 DR PDB: 1TBO; 14-OCT-96.
 DR PDB: 1TBR; 14-OCT-96.
 DR PDB: 1TOC; 23-JUL-97.
 DR PDB: 1VIT; 21-APR-97.
 DR PDB: 1YCP; 06-MAY-98.
 DR PDB: 1AOG; 17-JUN-98.
 DR PDB: 1AVG; 16-FEB-99.
 DR MEROPS; S01.217; .
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR002383; GLA blood.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003966; Prothrombin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR InterPro; IPR000294; VitK_dep_GLA.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR Pfam; PF00594; gla; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00001; GLABLOOD.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR01505; PROTHROMBIN.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00069; GLA; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; TYP-SPC; 1.
 DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
 KW Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
 KW Hydroxylase; Serine protease; Kringle; Signal; 3D-structure.
 FT SIGNAL 1 24
 FT PROPEP 25 43
 FT CHAIN 44 625
 FT PEPTIDE 44 199
 FT PEPTIDE 200 317
 FT CHAIN 318 366
 FT CHAIN 367 625
 FT DOMAIN 109 187
 FT DOMAIN 214 292
 FT SITE 367 625
 FT SITE 199 200
 FT SITE 317 318
 FT ACT_SITE 409 409
 FT ACT_SITE 465 465
 FT ACT_SITE 571 571
 FT MOD_RES 50 51
 FT MOD_RES 51 51
 FT MOD_RES 58 58
 FT MOD_RES 60 60
 FT MOD_RES 63 63
 FT MOD_RES 64 64
 FT MOD_RES 69 69
 FT MOD_RES 70 70
 FT MOD_RES 73 73

Query Match. 80.9%; Score 114; DB 1; Length 625;
 Best Local Similarity 95.2%; Pred. No. 2 4e-08;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 5 YKPGDKRGDACEGDSGGPFV 25
 ||| |||||||||

Db 556 YKPGDKRGDACEGDSGGPFV 576
 RESULT 3
 ID THRB_RAT STANDARD; PRT; 617 AA.
 AC P18292;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Prothrombin precursor (BC 3.4.21.5).
 GN F2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=90332426; PubMed=2377469;
 RA Dihanich M., Monard D.;
 RT "cDNA sequence of rat prothrombin.";
 RL Nucleic Acids Res. 18:4251-4251(1990).
 RN [2]
 RP SEQUENCE OF 383-617 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=92212913; PubMed=1557383;
 RA Banfield D.K., Macgillivray R.T.;
 RT "Partial characterization of vertebrate prothrombin cDNAs:
 RT amplification and sequence analysis of the B chain of thrombin from
 RT nine different species";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
 CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
 CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.
 CC -1- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOAL
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
 CC OF PROTHROMBIN TO THROMBIN.
 CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
 CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
 CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
 CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
 CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
 CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
 CC THROMBIN.
 CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
 CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
 CC BY FACTOR XA.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X52835; CAA37017.1; .
 CC EMBL; M81397; AAA42240.1; .
 CC PIR; S10511; S10511.
 CC HSSP; P00734; 1UVS.
 CC MEROPS; S01.217;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR002383; GLA_blood.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR003966; Prothrombin.

DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00051; kringle_2.
DR Pfam: PF00089; trypsin; 1.
DR Pfam: PF00594; gla; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00001; GLABLOOD.
DR PRINTS: PR00018; KRINGLE.
DR PRINTS: PR01505; PROTHROMBIN.
DR PRODOM: PD000395; kringle; 2.
DR SMART: SM00069; GLA; 1.
DR SMART: SM00130; KR; 2.
DR PROSITE: PS00020; TRYPSIN; 1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE: PS00021; KRINGLE_1; 2.
DR PROSITE: PS00070; KRINGLE_2; 2.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
KW Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
KW Hydrolase; Serine protease; Kringle; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 43
FT CHAIN 44 617
FT PEPTIDE 44 200
FT PEPTIDE 44 200
FT CHAIN 201 323
FT CHAIN 324 359
FT CHAIN 360 617
FT CHAIN 360 617
FT DOMAIN 109 187
FT DOMAIN 215 292
FT DOMAIN 360 617
FT SITE 200 201
FT SITE 323 324
FT SITE 359 360
FT SITE 402 402
FT ACT_SITE 458 458
FT ACT_SITE 564 564
FT MOD_RES 50 50
FT MOD_RES 51 51
FT MOD_RES 51 51
FT MOD_RES 58 58
FT MOD_RES 60 60
FT MOD_RES 63 63
FT MOD_RES 64 64
FT MOD_RES 69 69
FT MOD_RES 70 70
FT MOD_RES 73 73
FT MOD_RES 76 76
FT CARBOHYD 120 120
FT CARBOHYD 144 144
FT CARBOHYD 412 412
FT CARBOHYD 552 552
FT DISULFID 61 66
FT DISULFID 91 104
FT DISULFID 109 187
FT DISULFID 130 170
FT DISULFID 158 182
FT DISULFID 215 292
FT DISULFID 236 276
FT DISULFID 264 287
FT DISULFID 332 478
FT DISULFID 387 403
FT DISULFID 532 546
FT DISULFID 560 590
FT SEQUENCE 617 AA; 70411 MW; AD27D1B1445DB1D CRC64;
Query Match 65.2%; Score 92; DB 1; Length 617;
Best Local Similarity 76.2%; Pred. No. 2.5e-05;
Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 5 YKPDGGRGACGDSGGPFV 25
DB 549 FRYNDTRKGACGDSGGPFV 569

RESULT 4
ID THRB_MOUSE STANDARD; PRT; 618 AA.
AC P19221;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Prothrombin precursor (EC 3.4.21.5).
GN F2 OR CF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=91025551; PubMed=2222810;
RA Friesner Degen S.J., Schaffer L.A., Jamison C.S., Grant S.G.,
RA Fitzgibbon J.J., Pal J.A., Chapman V.M., Elliott R.W.;
RT "Characterization of the cDNA coding for mouse prothrombin and
RT localization of the gene on mouse chromosome 2.";
RL DNA Cell Biol. 9:487-498(1990).
RN [2]
RP SEQUENCE OF 384-618 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., Macgillivray R.T.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT amplification and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C,
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates
CC fibrinogen to fibrin and releases fibrinopeptide A and B.
CC -1- FM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL
CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
CC OF PROTHROMBIN TO THROMBIN.
CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES
CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
CC THROMBIN.
CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
CC BY FACTOR XA.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X52308; CAA36548.1;
CC EMBL: M81394; AAA0435.1;
CC PIR: A35827; A35827.
CC HSP: P00734; 1B7X.
CC MEROPS: S01.217;
CC MGD: MGI:88380; F2.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR002383; GLA_blood.

ID AC PRT: 461 AA.
 AC P04070; Q16001; Q15190; Q15189;
 DT 01-NOV-1986 (Rel. 03, Created)
 DT 01-NOV-1986 (Rel. 03, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)
 DE (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation
 DE factor XIV).
 GN PROC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=85270390; PubMed=2991887;
 RA Foster D.C., Yoshitake S., Davie E.W.;
 RT "The nucleotide sequence of the gene for human protein C";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4673-4677(1985).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=85265639; PubMed=2991859;
 RA Beckmann R.J., Schmidt R.J., Santerre R.F., Plutzky J., Crabtree G.R.,
 RA Long G.L.;
 RT "The structure and evolution of a 461 amino acid human protein C
 RT precursor and its messenger RNA, based upon the DNA sequence of
 RT cloned human liver cDNAs";
 RL Nucleic Acids Res. 13:5233-5247(1985).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=86120978; PubMed=3511471;
 RA Plutzky J., Hoskins J.A., Long G.L., Crabtree G.R.;
 RT "Evolution and organization of the human protein C gene";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:546-550(1986).
 RN [4]
 RN SEQUENCE FROM N.A.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE OF 106-461 FROM N.A.
 RX MEDLINE=84272714; PubMed=6589623;
 RA Foster D.C., Davie E.W.;
 RT "Characterization of a cDNA coding for human protein C";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4766-4770(1984).
 RN [6]
 RN CARBOHYDRATE-LINKAGE SITE ASN-371.
 RX MEDLINE=90293094; PubMed=1694179;
 RA Miletich J.P., Broze G.J. Jr.;
 RT "Beta protein C is not glycosylated at asparagine 329. The rate of
 RT translation may influence the frequency of usage at asparagine-X-
 RT cysteine sites";
 RL J. Biol. Chem. 265:11397-11404(1990).
 RN [7]
 RN HYDROXYLATION.
 RX MEDLINE=92184750; PubMed=1544894;
 RA Harris R.J., Ling V.T., Spellman M.W.;
 RT "O-linked fucose is present in the first epidermal growth factor
 RT domain of factor XII but not protein C";
 RL J. Biol. Chem. 267:5102-5107(1992).
 RN [8]
 RN 3D-STRUCTURE MODELING OF 175-450.
 RX MEDLINE=94272342; PubMed=8003977;
 RA Fisher C.L., Greengard J.S., Griffin J.H.;
 RT "Models of the serine protease domain of the human antithrombotic
 RT plasma factor activated protein C and its zymogen";
 RL Protein Sci. 3:588-599(1994).
 RN [9]
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 84-461.
 RX MEDLINE=97157472; PubMed=9003757;
 RA Mather T., Oganessyan V., Hof P., Huber R., Foundling S., Esmon C.,
 RA Bode W.;
 RT "The 2.8 A crystal structure of Gla-domainless activated protein C";
 RT

RL EMBO J. 15:6822-6831(1996).
 RN [10]
 RN REVIEW ON PROC VARIANTS.
 RX MEDLINE=93190290; PubMed=8446940;
 RA Reitsma P.H., Poort S.R., Bernardi F., Gandrille S., Long G.L.,
 RA Sala N., Cooper D.N.;
 RT "Protein C deficiency: a database of mutations. For the Protein C & S
 RT Subcommittee of the Scientific and Standardization Committee of the
 RT International Society on Thrombosis and Haemostasis";
 RL Thromb. Haemost. 69:77-84(1993).
 RN [11]
 RN VARIANT CYS-444.
 RX MEDLINE=87204221; PubMed=2437584;
 RA Romeo G., Hassan H.J., Staampfli S., Roncuzzi L., Cianetti L.,
 RA Leonardi A., Vicente V., Mannucci P.M., Bertina R.M., Peschle C.,
 RA Cortese R.;
 RT "Hereditary thrombophilia: identification of nonsense and missense
 RT mutations in the protein C gene";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2829-2832(1987).
 RN [12]
 RN VARIANT TRP-211 (LONDON-1).
 RX MEDLINE=90098906; PubMed=2602169;
 RA Grundy C.B., Chitollie A., Talbot S., Bevan D., Kakkar V.V.,
 RA Cooper D.N.;
 RT "Protein C London 1: recurrent mutation at Arg-169 (CGG-->TGG) in
 RT the protein C gene causing thrombosis";
 RL Nucleic Acids Res. 17:10513-10513(1989).
 RN [13]
 RN VARIANT CYS-272.
 RX MEDLINE=91329836; PubMed=1868249;
 RA Reitsma P.H., Poort S.R., Allaart C.F., Briet E., Bertina R.M.;
 RT "The spectrum of genetic defects in a panel of 40 Dutch families with
 RT symptomatic protein C deficiency type I: heterogeneity and founder
 RT effects";
 RL Blood 78:890-894(1991).
 RN [14]
 RN VARIANTS ALA-62 (VERMONT-1) AND MET-76.
 RX MEDLINE=92190481; PubMed=1347706;
 RA Bovill E.G., Tomczak J.A., Grant B., Bhushan F., Pillemer E.,
 RA Rainville I.R., Long G.L.;
 RT "Protein C Vermont: symptomatic type II protein C deficiency
 RT associated with two GLA domain mutations";
 RL Blood 79:1456-1465(1992).
 RN [15]
 RN VARIANT ASP-418 (HONG KONG-2).
 RX MEDLINE=92305321; PubMed=1611081;
 RA Sugahara Y., Miura O., Yuen P., Aoki N.;
 RT "Protein C deficiency Hong Kong 1 and 2: hereditary protein C
 RT deficiency caused by two mutant alleles, a 5-nucleotide deletion and
 RT a missense mutation";
 RL Blood 80:126-133(1992).
 RN [16]
 RN VARIANT LEU-289.
 RX MEDLINE=92380660; PubMed=1511988;
 RA Grundy C.B., Chisholm M., Kakkar V.V., Cooper D.N.;
 RT "A novel homozygous missense mutation in the protein C (PROC) gene
 RT causing recurrent venous thrombosis";
 RL Hum. Genet. 89:683-684(1992).
 RN [17]
 RN VARIANTS GLN-220 AND TRP-220.
 RX MEDLINE=92380661; PubMed=1511989;
 RA Grundy C.B., Schulman S., Tengborn L., Kakkar V.V., Cooper D.N.;
 RT "Two different missense mutations at Arg 178 of the protein C (PROC)
 RT gene causing recurrent venous thrombosis";
 RL Hum. Genet. 89:685-686(1992).
 RN [18]
 RN VARIANT GLN-220.
 RX MEDLINE=93250852; PubMed=1301959;
 RA Gandrille S., Vidaud M., Alach M., Alhenc-Gelas M., Fischer A.M.,
 RA Gouault-Heilman M., Toulon P., Flessinger J.N., Goossens M.;
 RT "Two novel mutations responsible for hereditary type I protein C
 RT deficiency: characterization by denaturing gradient gel
 RT electrophoresis";
 RT

RL Hum. Mutat. 1:491-500(1992).
 RN [19]
 RP VARIANT SER-334.
 RX MEDLINE-92276939; PubMed-1593215;
 RA Yamamoto K., Matsushita T., Sugitara I., Takamatsu J., Iwasaki E.,
 RA Wada H., Deguchi K., Shirakawa S., Saito H.,
 RT "Homozygous protein C deficiency: identification of a novel missense
 RT mutation that causes impaired secretion of the mutant protein C.";
 RL J. Lab. Clin. Med. 119:682-689(1992).
 RN [20]
 RP VARIANTS TRP-38; CYS-42; HIS-42; GLN-271 AND ASN-294.
 RX MEDLINE-93313192; PubMed-8324221;
 RA Gandrille S., Alenc-Gelas M., Gaussem P., Aillaud M.-F., Dupuy E.,
 RA Juhan-Vague I., Alach M.,
 RT "Five novel mutations located in exons III and IX of the protein C
 RT gene in patients presenting with defective protein C anticoagulant
 RT activity.";
 RL Blood 82:159-168(1993).
 RN [21]
 RP VARIANTS G-14; Q-211; Y-244; Q-253; L-321; C-328; I-385; T-388 AND
 RX MEDLINE-93371391; PubMed-8499565;
 RA Poort S.R., Pabinger-Fasching I., Mannhalter C., Reitsma P.H.,
 RA Bertina R.M.,
 RT "Twelve novel and two recurrent mutations in 14 Austrian families
 RT with hereditary protein C deficiency.";
 RL Blood Coagul. Fibrinolysis 4:273-280(1993).
 RN [22]
 RP VARIANT TRP-57.
 RX MEDLINE-93371396; PubMed-8499568;
 RA Millar D.S., Grundy C.B., Bignell P., Moffat E.H., Martin R.,
 RA Kaktar V.V., Cooper D.N.,
 RT "A Glu domain mutation (Arg 15-->Trp) in the protein C (PROC) gene
 RT causing type 2 protein C deficiency and recurrent venous
 RT thrombosis.";
 RL Blood Coagul. Fibrinolysis 4:345-347(1993).
 RN [23]
 RP VARIANTS R-145; L-210; W-211; T-243; L-321; M-340 AND Y-426.
 RX MEDLINE-94122329; PubMed-8292730;
 RA Tsay W., Greengard J.S., Montgomery R.R., McPherson R.A., Fucci J.C.,
 RA Koerber M.A., Coughlin J., Griffin J.H.,
 RT "Genetic mutations in ten unrelated American patients with
 RT symptomatic type I protein C deficiency.";
 RL Blood Coagul. Fibrinolysis 4:791-796(1993).
 RN [24]
 RP VARIANT SER-423.
 RX MEDLINE-94001606; PubMed-8398832;
 RA Marchetti G., Patracchini P., Gemmati D., Castaman G., Rodeghiero F.,
 RA Wacey A., Cooper D.N., Tuddenham E.G., Bernardi F.,
 RT "Symptomatic type II protein C deficiency caused by a missense
 RT mutation (Gly 381-->Ser) in the substrate-binding pocket.";
 RL Br. J. Haematol. 84:285-289(1993).
 RN [25]
 RP SEQUENCE OF 43-64 FROM N.A., AND VARIANT GLX-57 (YONAGO).
 Query Match 50.4%; Score 71; DB 1; Length 461;
 Best Local Similarity 81.2%; Pred. No. 0.015;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 10 GKRQDACEGDSGGPFV 25
 Db 392 GDRQDACEGDSGGPMV 407
 RESULT 7
 ID VDP_BOMMO STANDARD; PRT; 264 AA.
 AC Q07943;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vitellin-degrading protease precursor (EC 3.4.21.-) [Contains: Alpha-
 VTN protease; Beta-VTN protease].

OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 OC Ditrysia; Bombycoidea; Bombycidae; Bombyx.
 OX NCBI_TaxID-7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-N4; TISSUE-Egg;
 RX MEDLINE-92111263; PubMed-1764920;
 RA Ikeda M., Yaginuma T., Kobayashi M., Yamashita O.,
 RT "cDNA cloning, sequencing and temporal expression of the protease
 RT responsible for vitellin degradation in the silkworm, Bombyx mori.";
 RL Comp. Biochem. Physiol. 99B:405-411(1991).
 RN [2]
 RP PARTIAL SEQUENCE.
 RC STRAIN-Shunrei X Shogetu; TISSUE-Egg;
 RA Ikeda M., Sasaki T., Yamashita O.,
 RT "Purification and characterization of proteases responsible for
 RT vitellin degradation of the silkworm, Bombyx mori.";
 RL Insect Biochem. 20:725-734(1990).
 CC -|- FUNCTION: RESPONSIBLE FOR THE DEGRADATION OF VITELLIN IN EGGS AT
 CC THE HEAD PIGMENTATION STAGE.
 CC -|- DEVELOPMENTAL STAGE: APPEARED ON EGG DAY 7.5, BECAME MORE ACTIVE
 CC ON DAY 8-8.5 AND DISAPPEARED ON DAY 9.5.
 CC -|- PTM: CLEAVAGE AFTER ARG-27 LEADS TO BETA-VTN PROTEASE AND
 CC SUBSEQUENT CLEAVAGE AFTER ARG-89 LEADS TO ALPHA-VTN.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D16232; BAA03757.1; -;
 DR EMBL; D16233; BAA03758.1; -;
 DR PIR; S32794; S32794.
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.113; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser. protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Signal; Zymogen; Glycoprotein.
 FT SIGNAL 1 15 POTENTIAL.
 FT PROPEP 16 27 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 28 264 BETA-VTN PROTEASE.
 FT CHAIN 28 264 ALPHA-VTN PROTEASE CHAIN 1 (POTENTIAL).
 FT CHAIN 90 264 ALPHA-VTN PROTEASE CHAIN 2 (POTENTIAL).
 FT ACT_SITE 68 68 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 113 113 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 209 209 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT BINDING 203 203 SUBSTRATE (BY SIMILARITY).
 FT CARBOHYD 161 161 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 41 41 Y -> T (IN REF. 2).
 FT CONFLICT 56 56 T -> S (IN REF. 2).
 FT CONFLICT 94 94 Y -> T (IN REF. 2).
 SQ SEQUENCE 264 AA; 28521 MW; 6D536DD4184123AF CRC64;
 Query Match 50.0%; Score 70.5; DB 1; Length 264;
 Best Local Similarity 64.0%; Pred. No. 0.01;
 Matches 16; Conservative 2; Mismatches 2; Indels 5; Gaps 2;
 QY 1 ACTRYKPDGKRGDACEGDSGGPFV 25
 Db 195 ACT---PEGGK--DAQQDSDGGPLV 214

RN PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE-94289486; PubMed-8018728;
 RA Dhanach M.E., Spiess M.;
 RL "A novel serine proteinase-like sequence from human brain.";
 RT Biochim. Biophys. Acta 1218:225-228(1994).
 CC -1- FUNCTION: Protease whose physiological substrate is not yet known.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Highest expression in the thyroid gland. Also expressed in the prostate, salivary, and adrenal glands and in the colon testis and kidney.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@sib-sib.ch).
 CC EMBL; AF242195; AAG09469.1; -
 CC EMBL; AF242195; AAG09470.1; -
 CC EMBL; AF242195; AAG09471.1; -
 CC EMBL; AF242195; AAG09472.1; -
 CC EMBL; AF243527; AAG33354.1; -
 CC EMBL; X75363; CAA53145.1; ALT_SEQ.
 CC HSP; P00763; IDPO.
 CC MEROPS; S01.081; -
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; Tryp_SPC; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Glycoprotein; Signal; Zymogen;
 KW Alternative splicing.
 FT SIGNAL 1 16
 FT PROPEP 17 21
 FT CHAIN 22 256
 FT ACT_SITE 62 62
 FT ACT_SITE 106 106
 FT ACT_SITE 209 209
 FT CARBOHYD 171 171
 FT CARBOHYD 232 232
 FT VARSPIC 122 206
 FT VARSPIC 122 256
 FT VARSPIC 161 161
 FT VARSPIC 162 162
 FT CONFLICT 147 160
 FT SEQUENCE 256 AA; 28087 MW; B5EBF8D6022786B5 CRC64;
 Query Match 47.5%; Score 67; DB 1; Length 256;
 Best Local Similarity 64.7%; Pred. No. 0.029; Indels 3; Gaps 0;
 Matches 11; Conservative 3; Mismatches 3; Indels 3; Gaps 0;
 QY 9 EGKRGDACEGDSGGPFV 25
 ||: ::|||
 Db 198 EGRGAESCEGDSGGPLV 214
 RESULT 12
 ID HEPS_MOUSE STANDARD; PRT; 416 AA.
 AC O35453;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Serine protease hepsin (EC 3.4.21.-).
 GN HPN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 RN NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Liver;
 RX MEDLINE-98058912; PubMed-9395459;
 RA Vu T.-K.H., Liu R.W., Haakma C., Tomasek J.J., Howard E.W.;
 RT "Identification and cloning of the membrane-associated serine protease, hepsin, from mouse preimplantation embryos";
 RL J. Biol. Chem. 272:31315-31320(1997).
 CC -1- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CELL GROWTH AND MAINTENANCE OF CELL MORPHOLOGY.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@sib-sib.ch).
 CC EMBL; AF030065; AAB84221.1; -
 CC HSP; P00763; IDPO.
 CC MEROPS; S01.224; -
 CC MGD; MGI:1196620; Hpn.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; Tryp_SPC; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Transmembrane; Signal-anchor.
 KW CHAIN 1 161
 FT CHAIN 162 416
 FT DOMAIN 1 16
 FT DOMAIN 17 43
 FT DOMAIN 44 416
 FT DOMAIN 162 416
 FT ACT_SITE 202 202
 FT ACT_SITE 256 256
 FT ACT_SITE 352 352
 FT DISULFID 152 276
 FT DISULFID 187 203
 FT DISULFID 321 337
 FT DISULFID 348 380
 FT CARBOHYD 111 111
 FT SEQUENCE 416 AA; 44739 MW; 432194FF4004F848 CRC64;
 Query Match 47.5%; Score 67; DB 1; Length 416;
 Best Local Similarity 50.0%; Pred. No. 0.047; Indels 6; Gaps 1;
 Matches 15; Conservative 2; Mismatches 7; Indels 6; Gaps 1;
 QY 2 GTRYKPDGKRG-----DACEGDSGGPFV 25
 ||: |||
 Db 328 GNOIKPKMFCAGYPEGGIDACQDSGGPFV 357
 RESULT 13
 ID HEPS_HUMAN STANDARD; PRT; 417 AA.
 AC P05981;
 DT 01-NOV-1988 (Rel. 09, Created)


```

FT ACT_SITE 203 203 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 257 257 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 353 353 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 153 277 INTERCHAIN (BY SIMILARITY).
FT DISULFID 188 204 BY SIMILARITY.
FT DISULFID 322 338 BY SIMILARITY.
FT DISULFID 349 381 BY SIMILARITY.
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 417 AA; 45011 MW; B2086FF661E551D7 CRC64;

Query Match
Best Local Similarity 47.5%; Score 67; DB 1; Length 417;
Matches 15; Conservative 2; Mismatches 7; Indels 6; Gaps 1;

QY 2 GTRYKPDGKRG-----DACEGDSGGPFV 25
      | : || | | | | | | | | | | | | | |
DB 329 GNOIKPFCAGYPEGIDACQDSGGPFV 358

RESULT 14
KLKD.HUMAN
ID KLKD.HUMAN STANDARD; PRT; 277 AA.
AC Q9URK3; Q9Y433;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kallikrein 13 precursor (EC 3.4.21.-) (Kallikrein-like protein 4)
DE (KLK-L4).
DE KLK13 OR KLK14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20229789; PubMed=10766816;
RA Yousef G.M., Chang A., Diamandis E.P.;
RT Identification and characterization of KLK-L4, a new kallikrein-like
RT gene that appears to be down-regulated in breast cancer tissues.;
RL J. Biol. Chem. 275:11891-11898(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco M., Do L., Recala W., Terry A., Brower A., Ganes J.,
RA Danganan L., Ezler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Andreise I., Frankel M., Attix C., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT Sequence analysis of chromosome 19q13.4.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE OF 1-180 FROM N.A.
RP TISSUE=Uterus;
RA Ansgore W., Winkler U., Mexes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN PROSTATE, BREAST, TESTIS AND
CC SALIVARY GLAND.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. KALLIKREIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF135024; RAD26425.2; -
CC DR EMBL; AC011473; AAG23259.1; -
CC DR EMBL; AL050220; CAB43320.1; AUT INTT

```

```
DR HSSP; P00763; LDPO.
DR MEROPS; S01.306;
DR Genew; HGNC:6361; K1K13.
DR MIM; 605595;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; Signal.
FT SIGNAL 1 16
FT CHAIN 17 277
FT ACT_SITE 76 76
FT ACT_SITE 124 124
FT ACT_SITE 218 218
FT ACT_SITE 42 178
FT ACT_SITE 61 77
FT ACT_SITE 157 224
FT ACT_SITE 189 203
FT ACT_SITE 214 239
FT ACT_SITE 30 30
FT ACT_SITE 225 225
FT ACT_SITE 170 180
FT ACT_SITE 277 AA; 30570 MW; BA8A9E8DCF95D542 CRC64;
SQ SEQUENCE 277 AA; 30570 MW; BA8A9E8DCF95D542 CRC64;

Query Match 46.5%; Score 65.5; DB 1; Length 277;
Best Local Similarity 60.0%; Pred. No. 0.051;
Matches 15; Conservative 3; Mismatches 2; Indels 5; Gaps 2;

QY 1 AGTRYKDEGFRGACEDSGGPFV 25
DB 204 AGTR-----EGGK-DSCEGDSGGPLV 223

RESULT 15
ID PRTC_CANFA STANDARD; PRT; 157 AA.
AC Q28278;
DT 13-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vitamin-K dependent protein C (EC 3.4.21.69) (Autoprothrombin IIA)
DE (Anticoagulant protein C) (Blood coagulation factor XIV) (Fragment).
DE PROC.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94318474; PubMed=8043441;
RA Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;
RT "A comparative study of partial primary structures of the catalytic
RL region of mammalian protein C.";
RL Br. J. Haematol. 86:590-600(1994).
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va
CC and VIIa.
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; D43751; BAA07808.1;
DR HSSP; P04070; IPCU.
DR MEROPS; S01.218;
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Glycoprotein; Serine protease; Hydrolase.
FT NON_TER 1 1
FT ACT_SITE 26 26
FT ACT_SITE 125 125
FT ACT_SITE 96 110
FT ACT_SITE 121 149
FT ACT_SITE 17 17
FT ACT_SITE 78 78
FT ACT_SITE 157 157
FT ACT_SITE 157 AA; 17262 MW; E8B1BACF49220DFB CRC64;
SQ SEQUENCE 157 AA; 17262 MW; E8B1BACF49220DFB CRC64;

Query Match 46.1%; Score 65; DB 1; Length 157;
Best Local Similarity 75.0%; Pred. No. 0.035;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 10 GKRGDACEGDSGGPFV 25
DB 115 GDSRDACEGDSGGPMV 130

Search completed: May 12, 2003, 15:36:42
Job time : 15.5 secs
```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:32:37 ; Search time 56.5 Seconds
(without alignments)
91.171 Million cell updates/sec

Title: US-09-909-348-5
Perfect score: 141
Sequence: 1 ACTRYKPDGKRGDACEGDSGGPFV 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archesp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	83.0	235	6 Q28731	Q28731 oryctolagus
2	108	76.6	235	13 Q90387	Q90387 cynops pyrr
3	103	73.0	235	13 Q91004	Q91004 gecko gecko
4	103	73.0	607	13 Q91001	Q91001 gallus gall
5	103	73.0	608	13 Q9PTW7	Q9PTW7 struthio ca
6	99	70.2	239	13 Q91218	Q91218 oncorhynch
7	95	67.4	420	13 Q90504	Q90504 eptatretus
8	88	62.4	172	13 Q90FD1	Q90FD1 oncorhynch
9	82	58.2	234	13 Q90244	Q90244 acipenser t
10	70	49.6	378	5 Q8SY50	Q8SY50 drosophila
11	70	49.6	654	5 Q9VWU2	Q9VWU2 drosophila
12	70	49.6	974	13 Q90WD8	Q90WD8 bufo japoni
13	69	48.9	200	11 Q924U6	Q924U6 mus musculu
14	68	48.2	686	13 Q9DGC2	Q9DGC2 cyprinus ca
15	67.5	47.9	370	5 Q9VA44	Q9VA44 drosophila
16	67	47.5	255	4 Q96RQ0	Q96RQ0 homo sapien

17	67	47.5	258	5 Q9W508	Q9W508 drosophila
18	67	47.5	502	11 Q9CW97	Q9CW97 mus musculu
19	67	47.5	806	6 Q18783	Q18783 macropus eu
20	67	47.5	1524	13 Q91674	Q91674 xenopus lae
21	66	46.8	280	5 Q9V5X8	Q9V5X8 drosophila
22	66	46.8	562	6 Q8SQ23	Q8SQ23 sus scrofa
23	65	46.1	161	11 Q63109	Q63109 rattus norv
24	65	46.1	456	6 Q9TTR0	Q9TTR0 canis famil
25	65	46.1	481	11 Q54740	Q54740 mus musculu
26	65	46.1	481	11 Q99L32	Q99L32 mus musculu
27	65	46.1	481	11 Q88947	Q88947 mus musculu
28	65	46.1	482	11 Q63207	Q63207 rattus norv
29	64.5	45.7	161	11 Q60546	Q60546 mesocricetu
30	64.5	45.7	799	11 Q9DBI0	Q9DBI0 mus musculu
31	64.5	45.7	1801	5 Q8WSJ2	Q8WSJ2 bombyx mori
32	64	45.4	238	5 Q9WZC2	Q9WZC2 drosophila
33	64	45.4	253	4 Q8WZB4	Q8WZB4 homo sapien
34	64	45.4	254	5 Q76520	Q76520 stomoxys ca
35	64	45.4	259	5 Q9XV61	Q9XV61 ctenecephal
36	64	45.4	265	5 Q9XV55	Q9XV55 ctenecephal
37	64	45.4	309	5 Q27083	Q27083 tachyples
38	64	45.4	1004	13 P79953	P79953 xenopus lae
39	64	45.4	2616	5 Q9VFX5	Q9VFX5 drosophila
40	63.5	45.0	267	5 Q9BK47	Q9BK47 luidia foli
41	63	44.7	161	6 Q28511	Q28511 macaca mula
42	63	44.7	242	5 Q9XV59	Q9XV59 ctenecephal
43	63	44.7	255	5 Q9NBS9	Q9NBS9 glossina mo
44	63	44.7	267	5 Q9VLF5	Q9VLF5 drosophila
45	63	44.7	358	5 Q45029	Q45029 drosophila

ALIGNMENTS

RESULT 1

Q28731 ID Q28731 PRELIMINARY; PRT; 235 AA.
AC Q28731
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)
DE Thrombin (Fragment).
GN THROMBIN.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., MacGillivray R.T.A.;
RT *Partial characterization of vertebrate prothrombin cDNAs:
RT Amplification and sequence analysis of the B chain of thrombin from
RT nine different species.*
EL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC EMBL; M81396; AAA31477.1; -.
DR HSSP; P00734; 10VS.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser.protease_try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp.SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
FT NON_TER
SQ SEQUENCE 235 AA; 27093 MW; 92FF3E4F93B360E0 CRC64;

```

RX MEDLINE-92212913; PubMed-1557383;
RA Banfield D.K., MacGillivray R.T.A.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT Amplification and sequence analysis of the B chain of thrombin from
RT RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC
DR EMBL; M81392; AAA49309.1; -.
DR HSP; P00734; I87X.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; P00089; trypsin; 1.
DR PRINTS; P00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
FT NON_TER
FT
SQ SEQUENCE 235 AA; 26933 MW; 122A5C09F62276A CRC64;

Query Match 73.0%; Score 103; DB 13; Length 235;
Best Local Similarity 81.0%; Pred. No. 2.2e-07;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 YKDEGRKRGDACEGDSGGPFV 25
| : : : : : : : : : : : : : : :
Db 166 YSPEDSKRGDACEGDSGGPFV 186

RESULT 4
Q91001 PRELIMINARY; PRT; 607 AA.
ID Q91001
AC Q91001;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Thrombin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE-92212913; PubMed-1557383;
RA Banfield D.K., MacGillivray R.T.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT Amplification and sequence analysis of the B chain of thrombin from
RT RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE-94223694; PubMed-7513365;
RA Banfield D.K., Irwin D.M., Walz D.A., MacGillivray R.T.;
RT "Evolution of prothrombin; isolation and characterization of the cDNAs
RT encoding chicken and hagfish prothrombin.";
RL J. Mol. Evol. 38:177-187(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Banfield D.K.;
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
DR EMBL; M81391; AAA21619.1; -.
DR HSP; P00734; I0VS.
DR MEROPS; S01.217; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002383; GLA_blood.
DR InterPro; IPR000001; Kringle.

```

```
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.  
DR PROSITE; PS00021; KRINGLE_1; 2.  
DR PROSITE; PS00070; KRINGLE_2; 2.  
DR PROSITE; PS00240; TRYPSIN_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.  
DR PROSITE; PS00135; TRYPSIN_SER; 1.  
KW Hydrolase; Serine protease.  
SQ SEQUENCE 608 AA; 69392 MW; 11B974B9AE54EA2 CRC64;  
  
Query Match 73.0%; Score 103; DB 13; Length 608;  
Best Local Similarity 81.0%; Pred. No. 5.9e-07;  
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 YKPDGKRGDACEGDSGGPFV 25  
| |:: |||||  
DB 539 YSPEDSKRGDACEGDSGGPFV 559  
  
RESULT 6  
Q91218 PRELIMINARY; PRT; 239 AA.  
AC Q91218;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DI 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Thrombin (Fragment).  
GN THROMBIN.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
NCBI_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE=92212913; PubMed=1557383;  
RA Banfield D.K., MacGillivray R.T.A.;  
RT "Partial characterization of vertebrate prothrombin cDNAs:  
RT Amplification and sequence analysis of the B chain of thrombin from  
RT nine different species";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
CC EMBL; M81398; AAA49433.1; -.  
DR HSSP; P00734; 1B7X.  
DR MEROPS; S01_217; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR001254; Ser_protease_try.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.  
DR PROSITE; PS00135; TRYPSIN_SER; 1.  
KW Hydrolase; Serine protease.  
FT NON_TER 1  
SQ SEQUENCE 239 AA; 27396 MW; F0F43F9A3205BF38 CRC64;  
  
Query Match 70.2%; Score 99; DB 13; Length 239;  
Best Local Similarity 81.0%; Pred. No. 8.7e-07;  
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 5 YKPDGKRGDACEGDSGGPFV 25  
| |:: | |||||  
DB 166 FKPEQTGDACEGDSGGPFV 186  
  
RESULT 7  
Q90504 PRELIMINARY; PRT; 420 AA.  
AC Q90504;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DI 01-NOV-1996 (Tremblrel. 01, Last sequence update)
```

```

RN  SEQUENCE FROM N.A.
RA  Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;
RT  "Immune-relevant (including acute phase) genes identified in the
RT  livers of rainbow trout, Oncorhynchus mykiss, by means of suppression
RT  subtractive hybridization.";
RL  Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC  TRYPSIN FAMILY.
EMBL: AF281359; AAG30034.1; -.
DR  HSSP: P00734; 10VS.
DR  MEROPS: S01.217; -.
DR  InterPro: IPR001314; Chymotrypsin.
DR  InterPro: IPR001254; Ser.protease_Try.
DR  Pfam: PF00089; trypsin; 1.
DR  PRINTS: PR00722; CHYMOTRYPSIN.
DR  SMART: SM00020; Tryp_SPC; 1.
DR  PROSITE: PS0240; TRYPSIN_DOM; 1.
DR  PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW  Hydrolase; Serine protease.
FT  NON_TER 1
SQ  SEQUENCE 172 AA; 19873 MW; EC741FBA6957A887 CRC64;

Query Match 62.4%; Score 88; DB 13; Length 172;
Best Local Similarity 76.2%; Pred. No. 2.6e-05;
Matches 16; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 YKPDGKRGDACEGDSGGPFV 25
DB 99 FKPEQTGDAYEGDSGGPFV 119
:|:-| | | | | | | | | |
:|:-| | | | | | | | | |

RESULT 9
Q90244
ID Q90244 PRELIMINARY; PRT; 234 AA.
AC Q90244;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE Thrombin (Fragment).
GN THROMBIN.
OS Acipenser transmontanus (White sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OC NCBI_TaxID=7904;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., Macgillivray R.T.A.;
RT "Partial characterization of vertebrate prothrombin cDNAs:
RT Affiliation and sequence analysis of the B chain of thrombin from
RT nine different species.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
EMBL: M81399; AAA48514.1; -.
DR  HSSP: P00734; 2HNT.
DR  MEROPS: S01.217; -.
DR  InterPro: IPR001314; Chymotrypsin.
DR  InterPro: IPR001254; Ser.protease_Try.
DR  Pfam: PF00089; trypsin; 1.
DR  PRINTS: PR00722; CHYMOTRYPSIN.
DR  SMART: SM00020; Tryp_SPC; 1.
DR  PROSITE: PS0240; TRYPSIN_DOM; 1.
DR  PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR  PROSITE: PS00135; TRYPSIN_SER; 1.
KW  Hydrolase; Serine protease.
FT  NON_TER 1
SQ  SEQUENCE 234 AA; 26846 MW; 45C558D6618E0585 CRC64;

Query Match 58.2%; Score 82; DB 13; Length 234;

```

[illegible]

```

DR EMBL: AB070367; BAB63372.1; -.
DR MEROPS: S01.240; -.
DR InterPro: IPR000859; CUB_domain.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00431; CUB; 2.
DR Pfam: PF00089; trypsin; 2.
DR PROSITE: PS01180; CUB; 2.
DR PROSITE: PS02040; TRYPSIN_DOM; 2.
DR PROSITE: PS0134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 974 AA; 107647 MW; F19705A470465553 CRC64;

Query Match 49.6%; Score 70; DB 13; Length 974;
Best Local Similarity 68.4%; Pred. No. 0.076;
Matches 13; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

Qy 7 PDEGRGDACGDSGGPFV 25
I: || |||||
Db 227 PEGGR--DACQDSGGPFL 243

RESULT 13
Q924U6 PRELIMINARY; PRT; 200 AA.
AC Q924U6
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serine protease-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsui H., Takahashi T.;
RT "Mus musculus mRNA for serine protease-like gene spliced variant-1.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047758; BAB63919.1; -.
DR MEROPS: S01.106; -.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 200 AA; 22135 MW; 972959E4686EF603 CRC64;

Query Match 48.9%; Score 69; DB 11; Length 200;
Best Local Similarity 80.0%; Pred. No. 0.02;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 11 KRGDACGDSGGPFV 25
I: |||||
Db 125 KGGDACGDSGGPLV 139

RESULT 14
Q9DGC2 PRELIMINARY; PRT; 686 AA.
AC Q9DGC2
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cirs-A protein.
GN Cirs-A.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.

TISSUE=HEPATOPANCREAS;
MEDLINE=21115142; PubMed=11220628;
RA Nakao M., Ohsaka K., Kato Y., Fujiki K., Yano T.;
RT "Molecular cloning of Complement C1r/C1s/MASP2-like serine proteases
from the common carp (Cyprinus carpio).";
RL Immunogenetics 52:255-263(2001).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
DR EMBL: AB042609; BAB17845.1; -.
DR HSSP: P00763; IDPO.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000859; CUB_domain.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00431; CUB; 2.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00084; sushi; 2.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00042; CUB; 2.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00179; EGF_Ca; 1.
DR SMART: SM00020; TRYPSIN_DOM; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS01180; CUB; 2.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01187; EGF_Ca; UNKNOWN_1.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Hydroxylation;
Serine protease.
SQ SEQUENCE 686 AA; 76390 MW; 44AAD8F03FC67BCA CRC64;

Query Match 48.2%; Score 68; DB 13; Length 686;
Best Local Similarity 75.0%; Pred. No. 0.1;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 8 DEGRGDACGDSGGP 23
I: |||||
Db 618 DEVKRVDSQDGGSGP 633

RESULT 15
Q9VA44 PRELIMINARY; PRT; 370 AA.
AC Q9VA44
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG11313 protein.
GN CG11313.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.E.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Rogers J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Wortman J.R., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

```


RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AE003775; AAF57079.1; -.
DR HSSP; P00763; IDPO.
DR FlyBase; FBgn0039798; CG11313.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 370 AA; 40937 MW; E1E54E03F2CDE2AA CRC64;

Query Match 47.9%; Score 67.5; DB 5; Length 370;
Best Local Similarity 66.7%; Pred. No. 0.065;
Matches 12; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 9 EGG-RGDACGDSGGPFFV 25
II: |||||:|||||:
Db 309 EGRSGDSCDGGGGLM 326

Search completed: May 12, 2003, 15:38:46
Job time : 57.5 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:31:22 ; Search time 2.86 Seconds
(without alignments)
46.591 Million cell updates/sec

Title: US-09-909-348-4
Perfect score: 4
Sequence: 1 v 1

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4	100.0	2	22 AAG99717	ERA binding domain
2	4	100.0	2	22 AAG99741	ERA binding domain
3	4	100.0	2	22 AAB91888	Apoptosis related
4	4	100.0	3	10 AAP90664	New antihypertensi
5	4	100.0	3	10 AAP90665	New antihypertensi
6	4	100.0	3	10 AAP90668	New antihypertensi
7	4	100.0	3	10 AAP90678	New antihypertensi
8	4	100.0	3	12 AAP90543	Hypotensive oligop
9	4	100.0	3	13 AAR29095	Elastase-inhibiti
10	4	100.0	3	13 AAR29096	Cathepsin G inhibi

11	4	100.0	3	14 AAR34022	Metal trapping pep
12	4	100.0	3	14 AAR34023	Metal trapping pep
13	4	100.0	3	14 AAR34623	Chromophore-contg.
14	4	100.0	3	14 AAR38408	Elastase inhibitor
15	4	100.0	3	14 AAR38409	Elastase inhibitor
16	4	100.0	3	14 AAR38411	Cathepsin G inhibi
17	4	100.0	3	14 AAR59327	HTV protease inhib
18	4	100.0	3	15 AAR73934	Novel tripeptide i
19	4	100.0	3	15 AAR7853	Calpain inhibitor.
20	4	100.0	3	15 AAR48288	ACE inhibitor pep
21	4	100.0	3	15 AAR51439	IGF-1 analogue N-t
22	4	100.0	3	15 AAR58565	Angiotensin conver
23	4	100.0	3	15 AAR61094	ACE-inhibiting tri
24	4	100.0	3	16 AAR75205	Microbial emetic t
25	4	100.0	3	16 AAR82736	Apolipoprotein A-I
26	4	100.0	3	16 AAR82739	Apolipoprotein A-I
27	4	100.0	3	16 AAR82743	Apolipoprotein A-I
28	4	100.0	3	16 AAR82907	Non-RGD, non-VISGR
29	4	100.0	3	16 AAR71717	Potential cross-li
30	4	100.0	3	17 AAW02319	Beta-amyloid modul
31	4	100.0	3	18 AAW25187	LDV-peptide capabl
32	4	100.0	3	18 AAW25204	VTL-peptide capabl
33	4	100.0	3	18 AAW18872	Peptide for diagn
34	4	100.0	3	19 AAW76751	Murine signalling
35	4	100.0	3	19 AAW64738	Angiotensin II pep
36	4	100.0	3	19 AAW71120	Peptide AII(1-3) u
37	4	100.0	3	19 AAW56244	Anti-inflammatory
38	4	100.0	3	19 AAW56229	Anti-inflammatory
39	4	100.0	3	19 AAW56190	Anti-inflammatory
40	4	100.0	3	19 AAW50171	Signal-transducing
41	4	100.0	3	19 AAW52444	Loop region used i
42	4	100.0	3	19 AAW52444	Loop region used i
43	4	100.0	3	19 AAW52445	Loop region used i
44	4	100.0	3	19 AAW52449	Loop region used i
45	4	100.0	3	19 AAW52450	Loop region used i

ALIGNMENTS

RESULT 1
AAG99717
ID AAG99717 standard; Peptide; 2 AA.
XX AAG99717;
AC AAG99717;
DT 27-SEP-2001 (first entry)
XX ERA binding domain polypeptide SEQ ID NO 159.
DE ERA binding domain; Escherichia coli; GTPase; antimicrobial;
KW antibacterial; antibiotic; pathogenesis; infection; vaccine;
KW peptide therapy.
XX Streptococcus thermophilus.
OS
XX WO200153458-A2.
EN
XX 26-JUL-2001.
PD
XX 17-JAN-2001; 2001WO-US01786.
PF
XX 18-JAN-2000; 2000US-0176870.
PR
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
XX Lupas AN, Pearce KH;
PI WPI; 2001-476108/51.
DR
XX New ERA binding domain polypeptides and polynucleotides encoding them,
XX useful as research reagents and materials for discovery of treatments
PT

PT and diagnostics for diseases, or for genetic immunisation -
 PS Claim 1; Page 26; 279pp; English.

XX The present invention relates to ERA binding domain polypeptides
 CC (AAG99559-AAG99989 and AAM00010-AAM00021). The era gene in Escherichia
 CC coli codes for an essential GTPase protein able to autophosphorylate at
 CC serine and/or threonine residues. The protein has potential antimicrobial
 CC and antibacterial activity and is useful in screening for antagonists,
 CC agonists and for compounds with antibiotic activity. The proteins are
 CC also useful in determining their role in pathogenesis of infection,
 CC dysfunction and disease and could be used as part of a vaccine and/or
 CC peptide therapy.

XX Sequence 2 AA;

Query Match 100.0%; Score 4; DB 22; Length 2;
 Best Local Similarity 100.0%; Pred. No. 7.2e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 V 1
 DB 2 V 2

RESULT 2

AAG99741
 ID AAG99741 standard; Peptide; 2 AA.

XX AAG99741;

DT 27-SEP-2001 (first entry)

XX ERA binding domain polypeptide SEQ ID NO 183.

KW ERA binding domain; Escherichia coli; GTPase; antimicrobial;
 KW antibacterial; antibiotic; pathogenesis; infection; vaccine;
 KW peptide therapy.

OS Caenorhabditis elegans.

PN WO200153458-A2.

XX 26-JUL-2001.

PF 17-JAN-2001; 2001WO-US01786.

XX 18-JAN-2000; 2000US-0176870.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM PLC.

PI Lupas AN, Pearce KH;

XX WPI; 2001-476108/51.

XX New ERA binding domain polypeptides and polynucleotides encoding them,
 PT useful as research reagents and materials for discovery of treatments
 PT and diagnostics for diseases, or for genetic immunisation -

XX Claim 1; Page 29; 279pp; English.

XX The present invention relates to ERA binding domain polypeptides
 CC (AAG99559-AAG99989 and AAM00010-AAM00021). The era gene in Escherichia
 CC coli codes for an essential GTPase protein able to autophosphorylate at
 CC serine and/or threonine residues. The protein has potential antimicrobial
 CC and antibacterial activity and is useful in screening for antagonists,
 CC agonists and for compounds with antibiotic activity. The proteins are
 CC also useful in determining their role in pathogenesis of infection,
 CC dysfunction and disease and could be used as part of a vaccine and/or
 CC peptide therapy.

XX Sequence 2 AA;

Query Match 100.0%; Score 4; DB 22; Length 2;
 Best Local Similarity 100.0%; Pred. No. 7.2e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 V 1
 DB 1 V 1

RESULT 3

AAB91888

ID AAB91888 standard; Peptide; 2 AA.

XX AAB91888;

DT 22-JUN-2001 (first entry)

XX Apoptosis related peptide SEQ ID NO:1064.

KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimidyl; maleimido group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

OS Homo sapiens.

OS Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

PF 17-MAY-2000; 2000WO-US13576.

XX 17-MAY-1999; 99US-0134406.

XX 10-SEP-1999; 99US-0153406.

XX 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity

XX Disclosure; Page 543; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (iii) and a
 CC reactive group (ii) (e.g. succinimidyl and maleimido groups) attached to
 CC a less therapeutically active amino acid region (iv), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity
 CC in vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention.

XX Sequence 2 AA;

Query Match 100.0%; Score 4; DB 22; Length 2;
 Best Local Similarity 100.0%; Pred. No. 7.2e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 V 1
DB 1 V 1

RESULT 4

AAP90664
ID AAP90664 standard; protein; 3 AA.

XX AAP90664;

XX 26-MAY-1990 (first entry)

XX New antihypertensive peptide.

XX Antihypertensive peptide; angiotensin-converting enzyme;
KW ACE.

XX JPO1083096-A.

XX 28-MAR-1989.

XX 25-SEP-1987; 87JP-0241646.

XX 25-SEP-1987; 87JP-0241646.

XX (AJIN) AJINOMOTO KK.

XX WPI; 1989-136272/18.

XX New antihypertensive peptide(s) -
PT used to inhibit angiotensin-converting enzyme

XX ; ; 20pp; Japanese.

XX The peptide and its salts inhibit angiotensin-converting
CC enzyme (ACE) and are useful as antihypertensives. They may be
CC administered orally, parenterally or rectally in the form of tablets,
CC capsules, granules, powder, syrup, suspension, suppositories,
CC ointment, cream, gel, plaster, inhalation compen. or
CC injection at a dose of 0.001-1000, pref. 0.01-10, mg, 1-3 times per day.

XX SQ Sequence 3 AA;

Query Match 100.0%; Score 4; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 7.2e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 V 1
DB 2 V 2

RESULT 5

AAP90665
ID AAP90665 standard; protein; 3 AA.

XX AAP90665;

XX 26-MAY-1990 (first entry)

XX New antihypertensive peptide.

XX Antihypertensive peptide; angiotensin-converting enzyme;
KW ACE.

XX JPO1083096-A.

XX 28-MAR-1989.

XX 25-SEP-1987; 87JP-0241646.

XX 25-SEP-1987; 87JP-0241646.

XX

PA (AJIN) AJINOMOTO KK.

XX WPI; 1989-136272/18.

XX New antihypertensive peptide(s) -
PT used to inhibit angiotensin-converting enzyme

XX ; ; 20pp; Japanese.

XX The peptide and its salts inhibit angiotensin-converting
CC enzyme (ACE) and are useful as antihypertensives. They may be
CC administered orally, parenterally or rectally in the form of tablets,
CC capsules, granules, powder, syrup, suspension, suppositories,
CC ointment, cream, gel, plaster, inhalation compen. or
CC injection at a dose of 0.001-1000, pref. 0.01-10, mg, 1-3 times per day.

XX SQ Sequence 3 AA;

Query Match 100.0%; Score 4; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 7.2e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 V 1
DB 1 V 1

RESULT 6

AAP90668

ID AAP90668 standard; protein; 3 AA.

XX AAP90668;

XX 26-MAY-1990 (first entry)

XX New antihypertensive peptide.

XX Antihypertensive peptide; angiotensin-converting enzyme;
KW ACE.

XX JPO1083096-A.

XX 28-MAR-1989.

XX 25-SEP-1987; 87JP-0241646.

XX 25-SEP-1987; 87JP-0241646.

XX (AJIN) AJINOMOTO KK.

XX WPI; 1989-136272/18.

XX New antihypertensive peptide(s) -
PT used to inhibit angiotensin-converting enzyme

XX ; ; 20pp; Japanese.

XX The peptide and its salts inhibit angiotensin-converting
CC enzyme (ACE) and are useful as antihypertensives. They may be
CC administered orally, parenterally or rectally in the form of tablets,
CC capsules, granules, powder, syrup, suspension, suppositories,
CC ointment, cream, gel, plaster, inhalation compen. or
CC injection at a dose of 0.001-1000, pref. 0.01-10, mg, 1-3 times per day.

XX SQ Sequence 3 AA;

Query Match 100.0%; Score 4; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 7.2e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 V 1
DB 1 V 1

Db 1 V 1

RESULT 7
AAP90678
ID AAP90678 standard; protein; 3 AA.
XX AC AAP90678;
XX DT 26-MAY-1990 (first entry)
XX DE New antihypertensive peptide.
XX KW Antihypertensive peptide; angiotensin-converting enzyme;
XX KW ACE.
XX PN JP01083096-A.
XX PD 28-MAR-1989.
XX PF 25-SEP-1987; 87JP-0241646.
XX PR 25-SEP-1987; 87JP-0241646.
XX PA (AJIN) AJINOMOTO KK.
XX DR WPI; 1989-136272/18.
XX PT New antihypertensive peptide(s) -
XX PT used to inhibit angiotensin-converting enzyme
XX PS ; ; 20pp; Japanese.
XX CC The peptide and its salts inhibit angiotensin-converting
CC enzyme (ACE) and are useful as antihypertensives. They may be
CC administered orally, parenterally or rectally in the form of tablets,
CC capsules, granules, powder, syrup, suspension, suppositories,
CC ointment, cream, gel, plaster, inhalation compsn. or
CC injection at a dose of 0.001-1000, pref. 0.01-10, mg, 1-3 times per day.
XX SQ Sequence 3 AA;
Query Match 100.0%; Score 4; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. NO. 7.2e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 V 1
Db 1 V 1

RESULT 8
AAR10543
ID AAR10543 standard; Protein; 3 AA.
XX AC AAR10543;
XX DT 15-MAR-1991 (first entry)
XX DE Hypotensive oligopeptide.
XX KW Oral hypotensive; fig.
XX OS Ficus carica.
XX PN JP02282394-A.
XX PD 19-NOV-1990.
XX PF 24-APR-1989; 89JP-0104265.
XX PR 24-APR-1989; 89JP-0104265.
XX

PA (AGEN) AGENCY OF IND SCI TECH.
PA (SHOS) SHOWA SANGYO KK.
XX WPI; 1991-004480/01.
XX New oligopeptide hypotensive drug - based on alanine, valine,
PT asparagine, proline, isoleucine and arginine
XX Claim 1; Page 843; 9pp; Japanese.
XX Peptides may be derived from extract of fig, and are hypotensive
agents.
XX Sequence 3 AA;
Query Match 100.0%; Score 4; DB 12; Length 3;
Best Local Similarity 100.0%; Pred. NO. 7.2e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 V 1
Db 2 V 2

RESULT 9
AAR29095
ID AAR29095 standard; peptide; 3 AA.
XX AC AAR29095;
XX DT 19-APR-1993 (first entry)
XX DE Elastase-inhibiting fragment.
XX KW Cathepsin G; elastase; connective tissue; degradation; protease;
KW gout; rheumatoid arthritis; emphysema; ARDS;
XX adult respiratory distress syndrome; para-phenylene.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /note- "N-terminal amino acids of the peptides
FT of AAR29095-96 are linked by
FT -C(O)-para-phenylene-C(O)-"
FT Modified-site 4 /note- "C-terminal Val is in keto form, i.e. OH
FT replaced by CF2CF3"
XX PN WO9220357-A.
XX PD 26-NOV-1992.
XX PF 21-APR-1992; 92WO-US03288.
XX PR 23-MAY-1991; 91US-0704449.
XX PA (RICH) MERRELL DOW PHARM INC..
XX PI Angelastro MR, Bey P, Doherty NS, Janusz MJ, Mehdi S;
PI Peet NP;
XX DR WPI; 1992-415461/50.
XX PT New peptide derivs. used as cathepsin G and elastase inhibitors -
PT for treating gout, rheumatoid arthritis, inflammatory disorders,
PT emphysema and adult respiratory distress syndrome
XX Claim 18; Page 53; 53pp; English.
XX Inhibitors of cathepsin G and elastase for preventing connective
CC tissue degradation are chemically linked inhibitors of the proteases
CC elastase (AAR29095) and cathepsin G (AAR29096).

CC The N-terminal amino acids of the peptides of AAR29095-96 are
 CC linked by -C(O)-paraphenylene-C(O)-.
 CC The cpd. has an anti-inflammatory effect useful in the treatment
 CC of gout, rheumatoid arthritis and other inflammatory diseases and
 CC to prevent elastin mediated tissue damage. It can also be used
 CC in the treatment of emphysema and adult respiratory distress syndrome.
 XX

XX SQ Sequence 3 AA;

Query Match 100.0%; Score 4; DB 13; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.2e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
 Db 1 V 1

RESULT 10

AAR29096
 ID AAR29096 standard; peptide; 3 AA.

XX AC AAR29096;

XX DT 19-APR-1993 (first entry)

XX DE Cathepsin G inhibiting fragment.

XX KW Cathepsin G; elastase; connective tissue; degradation; protease;
 KW gout; rheumatoid arthritis; emphysema; ARDS;
 KW adult respiratory distress syndrome; para-phenylene.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Modified-site 1

FT /note= "N-terminal amino acids of the peptides
 of AAR29095-96 are linked by
 -C(O)-para-phenylene-C(O)-"

FT Modified-site 4

FT /note= "C-terminal Phe is in keto form, i.e. OH
 replaced by CF3"

XX WO9220357-A.

XX PN 26-NOV-1992.

XX PD 21-APR-1992; 92WO-US03288.

XX PF 23-MAY-1991; 91US-0704449.

XX PR (RICH) MERRELL DOW PHARM INC.

XX PA Angelastro MR, Bey P, Doherty NS, Janusz MJ, Mehdi S;
 PI Peet NP;

XX XX WPI; 1992-415461/50.

XX New peptide derivs. used as cathepsin G and elastase inhibitors -
 PT for treating gout, rheumatoid arthritis, inflammatory disorders,
 PT emphysema and adult respiratory distress syndrome

XX Claim 18; Page 53; 53pp; English.

XX Inhibitors of cathepsin G and elastase for preventing connective
 CC tissue degradation are chemically linked inhibitors of the proteases
 CC elastase (AAR29095) and cathepsin G (AAR29096).

CC The N-terminal amino acids of the peptides of AAR29095-96 are
 CC linked by -C(O)-paraphenylene-C(O)-.

CC The cpd. has an anti-inflammatory effect useful in the treatment
 CC of gout, rheumatoid arthritis and other inflammatory diseases and
 CC to prevent elastin mediated tissue damage. It can also be used
 CC in the treatment of emphysema and adult respiratory distress syndrome.

XX SQ Sequence 3 AA;

Query Match 100.0%; Score 4; DB 13; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.2e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
 Db 1 V 1

RESULT 11

AAR34022
 ID AAR34022 standard; peptide; 3 AA.

XX AC AAR34022;

XX DT 19-MAY-1993 (first entry)

XX DE Metal trapping peptide intermediate.

XX KW Electrostatic interaction; metalloprotein; binding site;
 KW protecting group.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "Boc protected"

FT Modified-site 2 /note= "Cys(Acm); Acm= acetamidomethyl-protected SH
 of cysteine"

FT Modified-site 3 /note= "OMe protected"

XX JP04346999-A.

XX PD 02-DEC-1992.

XX PF 24-MAY-1991; 91JP-0120196.

XX PR 24-MAY-1991; 91JP-0120196.

XX PA (ICHI-) ICHIKAWA GOSSEI KAGAKU KK.

XX DR WPI; 1993-021449/03.

XX New acyclic peptide with metal trapping activity - exhibits in
 PT vivo metal transporting action and can be used as drug,
 PT diagnostic agent or functional material

XX Claim 3; Page 2; 9pp; Japanese.

XX The acyclic peptide (see AAR31340) with metal trapping activity may be
 CC produced by fragment condensation using the peptide fragments given in
 CC AAR34022-29.

XX SQ Sequence 3 AA;

Query Match 100.0%; Score 4; DB 14; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.2e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
 Db 1 V 1

RESULT 12

AAR34023
 ID AAR34023 standard; peptide; 3 AA.

XX

AC AAR34023;
 XX
 DT 19-MAY-1993 (first entry)
 XX
 DE Metal trapping peptide intermediate.
 XX
 KW Electrostatic interaction; metalloprotein; binding site;
 KW protecting group.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "HCL-H-Val"
 FT Modified-site 2 /note= "Cys(Acm); Acm= acetamidomethyl-protected SH
 FT Modified-site 3 of cysteine"
 FT /note= "Ome protected"
 XX JP04346999-A.
 PN
 XX
 PD 02-DEC-1992.
 XX
 XX 24-MAY-1991; 91JP-0120196.
 XX
 XX 24-MAY-1991; 91JP-0120196.
 XX
 PA (TCHI-) ICHIKAWA GOSEI KAGAKU KK.
 XX
 XX WPI; 1993-021449/03.
 DR
 XX
 XX New acyclic peptide with metal trapping activity - exhibits in
 PT vivo metal transporting action and can be used as drug,
 PT diagnostic agent or functional material
 XX
 XX Claim 3; Page 2; 9pp; Japanese.
 XX
 XX The acyclic peptide (see AAR31340) with metal trapping activity may be
 CC produced by fragment condensation using the peptide fragments given in
 CC AAR34022-29.
 XX
 SQ Sequence 3 AA;
 Query Match 100.0%; Score 4; DB 14; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.2e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 V 1
 Db 1 V 1
 RESULT 13
 AAR34623
 ID AAR34623 standard; peptide; 3 AA.
 AC
 XX
 DT 12-NOV-1992 (first entry)
 XX
 DE Chromophore-contg. cpd. for determining ICE activity (3).
 XX
 KW Interleukin-lb; IL-lb; inflammation; immune; disease;
 KW diagnosis; 7-amino-4-methylcoumarin amide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-acetyl-tyrosinyl"
 FT Modified-site 3 /note= "alaninyl 7-amino-4-methylcoumarin amide"
 FT

XX EP528487-A.
 PN
 XX
 PD 24-FEB-1993.
 XX
 PD 08-AUG-1992; 92EP-0202450.
 PF
 XX 16-AUG-1991; 91US-0746455.
 PR
 PR 17-DEC-1991; 91US-0808994.
 XX
 XX (MERI) MERCK & CO INC.
 PA
 XX Chapman KT, Hagmann WK, Maccoss M, Mumford RA, Thornberry NA;
 PI Weidner JR;
 PI
 XX WPI; 1993-060350/08.
 DR
 XX New chromophore-contg. cpds. - for determining interleukin-lb
 PT convertase activity in diagnosis of inflammatory or immune-based
 PT disorders
 PT
 XX Claim 9; Page 42; 42pp; English.
 PS
 XX The cpd. is used for determining interleukin-lb convertase (ICE)
 CC activity. ICE has been implicated in inflammatory and immune-based
 CC diseases including diseases of the lungs and airways, CNS, eyes,
 CC ears, joints, bones and connective tissues, cardiovascular system
 CC (including the pericardium), GI and urogenital systems and skin and
 CC mucosal membranes.
 CC
 XX Sequence 3 AA;
 SQ
 Query Match 100.0%; Score 4; DB 14; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.2e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 V 1
 Db 2 V 2
 RESULT 14
 AAR38408
 ID AAR38408 standard; Protein; 3 AA.
 AC
 XX AAR38408;
 XX
 DT 29-OCT-1993 (first entry)
 DT
 XX Elastase inhibitor peptide #2.
 DE
 XX Inflammatory disease; prevention; rheumatoid arthritis; emphysema;
 KW neutrophil-mediated connective tissue degradation; gout; elastase;
 KW inhibition; adult respiratory distress syndrome.
 KW
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /label= N6-(2-carboxybenzoyl)lysine
 FT /note= "linked via a linking chain such as
 FT -CO-phenylene-CO- to the N-terminal of a
 FT Cathepsin G inhibiting molecule as described
 FT in AAR38410-11"
 FT Modified-site 3 /note= "May opt. be modified by COCOR, CF2CF3, CF3,
 FT CHF2, COOR3, CONHR3, CF2CHRC3ONHR, H, alkyl,
 FT aryl, aralkyl or COR; where R3 is H, alkyl,
 FT phenyl or benzyl and R is OH or alkoxy"
 FT
 XX ZA9203602-A.
 PN
 XX 24-FEB-1993.
 PD

```

XX 18-MAY-1992; 92ZA-0003602.
XX
XX 23-MAY-1991; 91US-0704499.
XX
XX (RICH ) MERRELL DOW PHARM INC.
XX
XX Angelastro MM, Bey P, Doherty NS, Janusz MJ, Mehdi S;
XX Peet NP;
XX
XX WPI; 1993-197380/24.
XX
XX New cathepsin G and elastase inhibitors - prevents connective
XX tissue degradation
XX
XX Claim 9-12; Page 51; 55pp; English.
XX
XX The sequences given in AAR38407-09 are elastase inhibiting peptides
XX which were produced by standard peptide synthesis methods. In the
XX context of the invention they are linked via their N-termini to a
XX Cathepsin G inhibiting peptide (see features table). The peptide
XX conjugates may be used to prevent neutrophil-mediated connective
XX tissue degradation associated with inflammatory diseases eg. gout
XX and rheumatoid arthritis. They may also be used for preventing
XX elastin-mediated tissue damage in the treatment of emphysema and
XX adult respiratory distress syndrome.
XX
XX Sequence 3 AA;
XX
XX Query Match 100.0%; Score 4; DB 14; Length 3;
XX Best Local Similarity 100.0%; Pred. No. 7.2e+05;
XX Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 V 1
XX 3 V 3
XX
XX RESULT 15
XX AAR38409
XX ID AAR38409 standard; Protein; 3 AA.
XX
XX AC AAR38409;
XX
XX DT 29-OCT-1993 (first entry)
XX
XX DE Elastase inhibitor peptide #3.
XX
XX KW Inflammatory disease; prevention; rheumatoid arthritis; emphysema;
XX neutrophil-mediated connective tissue degradation; gout; elastase;
XX inhibition; adult respiratory distress syndrome.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX Modified-site 1
XX /note- "Linked via a linking chain such as
XX -CO-phenylene-CO- to the N-terminal of a
XX Cathepsin G inhibiting molecule as described
XX in AAR38410-11"
XX
XX Modified-site 3
XX /note- "May opt. be modified by COOR, CF2CF3, CF3,
XX CHF2, COOR3, CONHR3, CF2CHR3CONHR, H, alkyl,
XX aryl, aralkyl or COR; where R3 is H, alkyl,
XX phenyl or benzyl and R is OH or alkoxy"
XX
XX PN ZA9203602-A.
XX
XX XX
XX PD 24-FEB-1993.
XX
XX PF 18-MAY-1992; 92ZA-0003602.
XX
XX PR 23-MAY-1991; 91US-0704499.

```

```

XX (RICH ) MERRELL DOW PHARM INC.
XX
XX Angelastro MM, Bey P, Doherty NS, Janusz MJ, Mehdi S;
XX Peet NP;
XX
XX WPI; 1993-197380/24.
XX
XX New cathepsin G and elastase inhibitors - prevents connective
XX tissue degradation
XX
XX Claim 9-12; Page 51; 55pp; English.
XX
XX The sequences given in AAR38407-09 are elastase inhibiting peptides
XX which were produced by standard peptide synthesis methods. In the
XX context of the invention they are linked via their N-termini to a
XX Cathepsin G inhibiting peptide (see features table). The peptide
XX conjugates may be used to prevent neutrophil-mediated connective
XX tissue degradation associated with inflammatory diseases eg. gout
XX and rheumatoid arthritis. They may also be used for preventing
XX elastin-mediated tissue damage in the treatment of emphysema and
XX adult respiratory distress syndrome.
XX
XX Sequence 3 AA;
XX
XX Query Match 100.0%; Score 4; DB 14; Length 3;
XX Best Local Similarity 100.0%; Pred. No. 7.2e+05;
XX Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 V 1
XX 1 V 1
XX
XX Search completed: May 12, 2003, 15:36:05
XX Job time : 3.86 secs

```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:33:37 ; Search time 0.96 Seconds
(without alignments)
30.649 Million cell updates/sec

Title: US-09-909-348-4

Perfect score: 4

Sequence: 1 V 1

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	100.0	2	1	US-08-122-510-9
2	4	100.0	2	1	Sequence 9, Appli
3	4	100.0	2	1	Sequence 10, Appl
4	4	100.0	2	1	Sequence 14, Appl
5	4	100.0	2	1	Sequence 5, Appli
6	4	100.0	2	2	Sequence 18, Appl
7	4	100.0	2	4	Sequence 29, Appl
8	4	100.0	2	4	Sequence 30, Appl
9	4	100.0	3	1	PCT-US94-10475-5
10	4	100.0	3	1	US-07-729-099-5
11	4	100.0	3	1	US-07-791-213D-38
12	4	100.0	3	1	US-08-118-135A-8
13	4	100.0	3	1	US-07-947-035-3
14	4	100.0	3	1	US-08-384-618-1
15	4	100.0	3	1	US-07-729-353-6
16	4	100.0	3	1	US-08-079-812-139
17	4	100.0	3	1	US-08-122-510-3
18	4	100.0	3	1	US-08-122-510-7
19	4	100.0	3	1	US-08-122-510-8
20	4	100.0	3	1	US-08-122-510-11
21	4	100.0	3	1	US-08-122-510-12
22	4	100.0	3	1	US-08-122-510-13
23	4	100.0	3	1	US-08-037-486-11
24	4	100.0	3	1	US-08-127-904-6
25	4	100.0	3	1	US-08-127-904-7
26	4	100.0	3	1	US-08-221-582A-1
27	4	100.0	3	1	US-08-221-582A-2
					Sequence 9, Appli

28 4 100.0 3 1 US-08-213-897A-10 Sequence 10, Appli
29 4 100.0 3 1 US-08-354-240A-16 Sequence 16, Appli
30 4 100.0 3 1 US-08-321-585A-1 Sequence 1, Appli
31 4 100.0 3 1 US-08-257-392-5 Sequence 5, Appli
32 4 100.0 3 1 US-08-634-060-36 Sequence 36, Appli
33 4 100.0 3 1 US-08-338-282-11 Sequence 11, Appli
34 4 100.0 3 1 US-07-955-905A-11 Sequence 11, Appli
35 4 100.0 3 1 US-08-233-150A-38 Sequence 38, Appli
36 4 100.0 3 1 US-08-408-604A-164 Sequence 164, App
37 4 100.0 3 2 US-08-478-386A-40 Sequence 40, Appl
38 4 100.0 3 2 US-08-292-597-40 Sequence 40, Appl
39 4 100.0 3 2 US-08-408-858A-14 Sequence 14, Appl
40 4 100.0 3 2 US-08-408-858A-17 Sequence 17, Appl
41 4 100.0 3 2 US-08-408-858A-21 Sequence 21, Appl
42 4 100.0 3 2 US-08-612-785B-19 Sequence 19, Appl
43 4 100.0 3 2 US-08-451-822A-18 Sequence 18, Appl
44 4 100.0 3 2 US-08-388-653-40 Sequence 40, Appl
45 4 100.0 3 2 US-08-473-985-40 Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-08-122-510-9
; Sequence 9, Application US/08122510
; Patent No. 5464821
; GENERAL INFORMATION:
; APPLICANT: AASMUL-OLSEN, Stig
; APPLICANT: WIDMER, Fred
; APPLICANT: GAURI, Kallash
; TITLE OF INVENTION: Small peptidic compounds useful for the
; TITLE OF INVENTION: treatment of Glaucoma
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Dr.
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/122,510
; FILING DATE: 24-SEP-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0532/91
; FILING DATE: 25-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK92/00095
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Iwanicki, John P
; REGISTRATION NUMBER: 34,628
; REFERENCE/DOCKET NUMBER: 93,848
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910/221-5317
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide

;
; LOCATION: 1..2
; OTHER INFORMATION: /note= "where x is Benzylester
; OTHER INFORMATION:
US-08-122-510-9

Query Match 100.0%; Score 4; DB 1; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
Db 2 V 2

RESULT 2
US-08-122-510-10
; Sequence 10, Application US/08122510
; Patent No. 5464821
; GENERAL INFORMATION:
; APPLICANT: AASMUL-OLSEN, Stig
; APPLICANT: WIDMER, Fred
; APPLICANT: GAURI, Kailash
; TITLE OF INVENTION: Small peptidic compounds useful for the
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Dr.
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/122,510
; FILING DATE: 24-SEP-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0532/91
; FILING DATE: 25-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK92/00095
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Iwanicki, John P.
; REGISTRATION NUMBER: 34,628
; REFERENCE/DOCKET NUMBER: 93,848
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910/221-5317
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..2
; OTHER INFORMATION: /note= "where x is Benzylester
; OTHER INFORMATION:
US-08-122-510-10

Query Match 100.0%; Score 4; DB 1; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
Db 2 V 2

RESULT 3
US-08-122-510-14
; Sequence 14, Application US/08122510
; Patent No. 5464821
; GENERAL INFORMATION:
; APPLICANT: AASMUL-OLSEN, Stig
; APPLICANT: WIDMER, Fred
; APPLICANT: GAURI, Kailash
; TITLE OF INVENTION: Small peptidic compounds useful for the
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Dr.
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/122,510
; FILING DATE: 24-SEP-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0532/91
; FILING DATE: 25-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK92/00095
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Iwanicki, John P.
; REGISTRATION NUMBER: 34,628
; REFERENCE/DOCKET NUMBER: 93,848
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910/221-5317
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..2
; OTHER INFORMATION: /note= "Boc-Gly Val-Obzl"
; OTHER INFORMATION:
US-08-122-510-14

Query Match 100.0%; Score 4; DB 1; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
Db 2 V 2

RESULT 4
US-08-127-904-5
; Sequence 5, Application US/08127904
; Patent No. 5470951
; GENERAL INFORMATION:
US-08-127-904-5

QY 1 V 1
Db 2 V 2

RESULT 3
US-08-122-510-14
; Sequence 14, Application US/08122510
; Patent No. 5464821
; GENERAL INFORMATION:
; APPLICANT: AASMUL-OLSEN, Stig
; APPLICANT: WIDMER, Fred
; APPLICANT: GAURI, Kailash
; TITLE OF INVENTION: Small peptidic compounds useful for the
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 S. Wacker Dr.
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/122,510
; FILING DATE: 24-SEP-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 0532/91
; FILING DATE: 25-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK92/00095
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Iwanicki, John P.
; REGISTRATION NUMBER: 34,628
; REFERENCE/DOCKET NUMBER: 93,848
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910/221-5317
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..2
; OTHER INFORMATION: /note= "Boc-Gly Val-Obzl"
; OTHER INFORMATION:
US-08-122-510-14

Query Match 100.0%; Score 4; DB 1; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
Db 2 V 2

RESULT 4
US-08-127-904-5
; Sequence 5, Application US/08127904
; Patent No. 5470951
; GENERAL INFORMATION:
US-08-127-904-5

```

; APPLICANT: Eugene Roberts
; TITLE OF INVENTION: Method For Antagonizing
; TITLE OF INVENTION: Amnestic Effects of Amyloid n
; TITLE OF INVENTION: Protein and Improving the
; TITLE OF INVENTION: Quality of Life in Individuals
; TITLE OF INVENTION: With Alzheimer Disease
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: City of Hope
; STREET: 1500 East Duarte Road
; CITY: Duarte
; STATE: California
; COUNTRY: United States of America
; ZIP: 91010-0269
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3M Double Density 5 1/4" diskette
; COMPUTER: Wang PC
; OPERATING SYSTEM: MS DOS Version 3.20
; SOFTWARE: Microsoft
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/127,904
; FILING DATE: 29 September 1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: No. 5470951e
; ATTORNEY/AGENT INFORMATION:
; NAME: Irons, Edward S.
; REGISTRATION NUMBER: 16,541
; REFERENCE/DOCKET NUMBER: No. 5470951e
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 783-6040
; TELEFAX: (202) 783-6031
; TELEX: No. 5470951e
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2
; TYPE: Amino Acid
; STRANDEDNESS:
; TOPOLOGY: Unknown
; US-08-127-904-5

```

```

Query Match 100.0%; Score 4; DB 1; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 V 1
Db 1 V 1

```

```

; RESULT 5
; US-08-603-737-18
; Sequence 18, Application US/08603737
; Patent No. 5849691
; GENERAL INFORMATION:
; APPLICANT: Majer, et al.
; TITLE OF INVENTION: Peptidomimetic Inhibitors
; TITLE OF INVENTION: of Cathepsin D and Plasmeppins I and II
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/603,737
; FILING DATE: February 20, 1996

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth A. Bogosian
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 2026-4186
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; DESCRIPTION: No
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE: C-terminal
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: Xaa is
; OTHER INFORMATION: phenylstatine. CH3(CH2)3CO is at the
; OTHER INFORMATION: amino terminus, and CH3(CH2)4NH is at
; OTHER INFORMATION: the carboxy terminus.
; US-08-603-737-18

```

```

Query Match 100.0%; Score 4; DB 2; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 V 1
Db 2 V 2

```

```

; RESULT 6
; US-09-264-709A-29
; Sequence 29, Application US/09264709A
; Patent No. 6320024
; GENERAL INFORMATION:
; APPLICANT: Roberts, Eugene
; TITLE OF INVENTION: Method for Design of Substances that Enhance Memory and
; TITLE OF INVENTION: Improve the Quality of Life
; FILE REFERENCE: 2124-310
; CURRENT APPLICATION NUMBER: US/09/264,709A
; CURRENT FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 08/797,782
; PRIOR FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 2
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: memory-modulating peptide
; US-09-264-709A-29

```

```

Query Match 100.0%; Score 4; DB 4; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 V 1
Db 1 V 1

```

RESULT 7
US-09-264-709A-30
; Sequence 30, Application US/09264709A
; Patent No. 6320024
; GENERAL INFORMATION:
; APPLICANT: Roberts, Eugene
; TITLE OF INVENTION: Method for Design of Substances that Enhance Memory and
; FILE OF INVENTION: Improve the Quality of Life
; FILE REFERENCE: 2124-310
; CURRENT APPLICATION NUMBER: US/09/264,709A
; CURRENT FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 08/797,782
; PRIOR FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 2
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: memory-modulating peptide
US-09-264-709A-30

Query Match 100.0%; Score 4; DB 4; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
DB 2 V 2

RESULT 8
PCT-US94-10475-5
; Sequence 5, Application PC/TUS9410475
; GENERAL INFORMATION:
; APPLICANT: Eugene Roberts
; TITLE OF INVENTION: Method For
; TITLE OF INVENTION: Antagonizing Amnestic
; TITLE OF INVENTION: Effects of Amyloid n
; TITLE OF INVENTION: Protein and Improving
; TITLE OF INVENTION: the Quality of Life
; TITLE OF INVENTION: in Individuals
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: City of Hope
; STREET: 1500 East Duarte Road
; CITY: Duarte
; STATE: California
; COUNTRY: United States of America
; ZIP: 91010-0269
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3M Double Density 5 1/4"
; MEDIUM TYPE: diskette
; COMPUTER: Wang PC
; OPERATING SYSTEM: MS DOS Version 3.20
; SOFTWARE: Microsoft
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10475
; FILING DATE: 16 September 1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA: U. S. Application
; PRIOR APPLICATION DATA: Serial No.
; PRIOR APPLICATION DATA: 08/127,904; filed
; PRIOR APPLICATION DATA: 29 September 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Irons, Edward S.
; REGISTRATION NUMBER: 16,541
; REFERENCE/DOCKET NUMBER: None

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 626-3564 or 783-6030
; TELEFAX: (202) 783-6031
; TELEX: None
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2
; TYPE: Amino Acid
; STRANDEDNESS:
; TOPOLOGY: Unknown
PCT-US94-10475-5

Query Match 100.0%; Score 4; DB 5; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
DB 1 V 1

RESULT 9
US-07-729-099-5
; Sequence 5, Application US/07729099
; Patent No. 5403581
; GENERAL INFORMATION:
; APPLICANT: Binger, Mary-Helen
; APPLICANT: Pasamontes, Luis
; TITLE OF INVENTION: Coccidiosis Vaccines
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.
; ZIP: 07110

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/729,099
; FILING DATE: 19910712
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Elmeria tenella
US-07-729-099-5

Query Match 100.0%; Score 4; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
DB 3 V 3

RESULT 10
 US-07-791-213D-38
 ; Sequence 38, Application US/07791213D
 ; Patent No. 5409895
 ; GENERAL INFORMATION:
 ; APPLICANT: MORISHITA, Hiideaki
 ; APPLICANT: KANAMORI, Toshinori
 ; APPLICANT: NOBUHARA, Masahiro
 ; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
 ; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
 ; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
 ; TITLE OF INVENTION: TREATING USING THE SAME
 ; NUMBER OF SEQUENCES: 108
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis
 ; STREET: P.O. Box 1404
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/791,213D
 ; FILING DATE: 13-NOV-1991
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-306745
 ; FILING DATE: 13-NOV-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Neuth, Donna M
 ; REGISTRATION NUMBER: 36,607
 ; REFERENCE/DOCKET NUMBER: 029650-032
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 836-6620
 ; TELEFAX: (703) 836-2021
 ; INFORMATION FOR SEQ ID NO: 38:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-07-791-213D-38
 Query Match 100.0%; Score 4; DB 1; Length 3;
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 V 1
 Db 3 V 3
 RESULT 11
 US-08-118-135A-8
 ; Sequence 8, Application US/08118135A
 ; Patent No. 5437982
 ; GENERAL INFORMATION:
 ; APPLICANT: Catterall, W.A. and Eaholtz, G.
 ; TITLE OF INVENTION: "Specific Inactivation Gate Inhibitors of the Sodium Channel"
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
 ; STREET: 2800 Pacific First Center, 1420 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98101-2347
 ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette-5.25 inch, 1.2Mb storage
 ; COMPUTER: IBM PC/386 Compatible
 ; OPERATING SYSTEM: MS-DOS 4.01
 ; SOFTWARE: Word 5.5-t
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/118,135A
 ; FILING DATE: September 8, 1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA: No. 5437982e
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brogerick, Thomas F.
 ; REGISTRATION NUMBER: 31,332
 ; REFERENCE/DOCKET NUMBER: UOFW16993
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 1-206-682-8100; 1-206-224-0709 (direct)
 ; TELEFAX: 1-206-224-0779
 ; TELEX: 4938023
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; DESCRIPTION: page 22, line 17; VFM
 ; US-08-118-135A-8
 Query Match 100.0%; Score 4; DB 1; Length 3;
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 V 1
 Db 1 V 1
 RESULT 12
 US-07-947-035-3
 ; Sequence 3, Application US/07947035
 ; Patent No. 5444045
 ; GENERAL INFORMATION:
 ; APPLICANT: Francis, Geoffrey L.
 ; APPLICANT: Walton, Paul E.
 ; APPLICANT: Ballard, Francis J.
 ; APPLICANT: McMurty, John P.
 ; APPLICANT: Phelps, Patricia V.
 ; TITLE OF INVENTION: Method of Administering IGF-1, IGF-2,
 ; TITLE OF INVENTION: and Analogs Thereof to Birds
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kenneth D. Sibley
 ; STREET: P.O. Drawer 34009
 ; CITY: Charlotte
 ; STATE: No. 5444045th Carolina
 ; COUNTRY: US
 ; ZIP: 28234
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/947,035
 ; FILING DATE: 17-SEP-1992
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sibley, Kenneth D.
 ; REGISTRATION NUMBER: 31,665
 ; REFERENCE/DOCKET NUMBER: 5175-59
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (919) 881-3140
 ; TELEFAX: (919) 881-3175
 ; TELEX: 575102
 ; INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 3 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO

US-07-947-035-3

Query Match 100.0%; Score 4; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1

Db 1 V 1

RESULT 13

US-08-384-618-1
Sequence 1, Application US/08384618
Patent No. 5449661
GENERAL INFORMATION:
APPLICANT: Nakamura, Yasunori
TITLE OF INVENTION: Angiotensin Converting Enzyme Inhibitor
TITLE OF INVENTION: and Method for Preparing Same
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby
STREET: 805 Third Ave.
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,618
FILING DATE:
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/090,530
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, S. Peter

REGISTRATION NUMBER: 25,351
REFERENCE/DOCKET NUMBER: 4703/08827
TELEPHONE: (212)527-7700
TELEFAX: (212)753-7513
TELEX: 236687

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO

US-08-384-618-1
Query Match 100.0%; Score 4; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1

Db 1 V 1

RESULT 14

US-07-729-353-6
Sequence 6, Application US/07729353
Patent No. 5464819
GENERAL INFORMATION:
APPLICANT: Suzuki, Naoyoshi
TITLE OF INVENTION: A No. 5464819el Physiologically Active Peptide Having
TITLE OF INVENTION: Immunoregulatory Activities
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 301 N. Washington Street
CITY: Falls Church
STATE: Virginia
COUNTRY: United States of America
ZIP: 22040-0747

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/729,353
FILING DATE: 19910712
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-182714
FILING DATE: 12-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Svensson, Leonard R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 216-273P
TELEPHONE: 703 241 1300
TELEFAX: 703 532 3407
TELEX: 248345

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-07-729-353-6
Query Match 100.0%; Score 4; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1

Db 2 V 2

RESULT 15

US-08-079-812-139
Sequence 139, Application US/08079812
Patent No. 5464820
GENERAL INFORMATION:
APPLICANT: BURTON, James
APPLICANT: DONG, Zhengxin
APPLICANT: FRIGO, Timothy B.
TITLE OF INVENTION: SPECIFIC INHIBITORS OF TISSUE KALLIKREIN
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 555 13th Street, N.W., Suite 500 East
CITY: Washington
STATE: D.C.

COUNTRY: U.S.A.
ZIP: 20004-1109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/079,812
FILING DATE: 22-JUN-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Remenick, James
REGISTRATION NUMBER: 36,902
REFERENCE/DOCKET NUMBER: 16865-0118
TELEPHONE: (202) 639-7700
TELEFAX: (202) 639-7832
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 3 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
US-08-079-812-139

Query Match 100.0%; Score 4; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 V 1
Db 1 V 1

Search completed: May 12, 2003, 15:40:40
Job time : 1.96 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:38:52 ; Search time 1.94 Seconds
(without alignments)
47.436 Million cell updates/sec

Title: US-09-909-348-4

Perfect score: 4

Sequence: 1 V 1

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	100.0	2	10	US-09-982-172-145
2	4	100.0	2	10	US-09-982-172-152
3	4	100.0	3	8	US-08-424-5508-186
4	4	100.0	3	9	US-08-826-290-464
5	4	100.0	3	9	US-09-900-936-11
6	4	100.0	3	9	US-09-828-272A-1
7	4	100.0	3	9	US-10-074-956-5
8	4	100.0	3	9	US-10-039-876A-6
9	4	100.0	3	9	US-10-015-979-63
10	4	100.0	3	9	US-10-015-979-71
11	4	100.0	3	10	US-09-803-854-9
12	4	100.0	3	10	US-09-804-866-14
13	4	100.0	3	10	US-09-771-192-11
14	4	100.0	3	10	US-09-853-918-6
15	4	100.0	3	10	US-09-850-061A-31
16	4	100.0	3	10	US-09-850-061A-32
17	4	100.0	3	10	US-09-850-061A-33
18	4	100.0	3	10	US-09-850-061A-37
19	4	100.0	3	10	US-09-850-061A-38

20 4 100.0 3 10 US-09-972-475-19
21 4 100.0 3 10 US-09-895-443-10
22 4 100.0 3 10 US-09-984-056-25
23 4 100.0 3 10 US-09-982-172-47
24 4 100.0 3 10 US-09-982-172-58
25 4 100.0 3 10 US-09-982-172-84
26 4 100.0 3 10 US-09-837-697A-11
27 4 100.0 3 10 US-09-984-057-25
28 4 100.0 4 1 US-08-464-363-25
29 4 100.0 4 8 US-08-859-699-17
30 4 100.0 4 8 US-08-839-248-1
31 4 100.0 4 8 US-08-839-248-2
32 4 100.0 4 8 US-08-839-248-3
33 4 100.0 4 8 US-08-839-248-4
34 4 100.0 4 8 US-08-450-842-77
35 4 100.0 4 8 US-08-450-842-92
36 4 100.0 4 8 US-08-987-689A-15
37 4 100.0 4 8 US-08-987-689A-16
38 4 100.0 4 8 US-08-808-031A-52
39 4 100.0 4 8 US-08-681-219-10
40 4 100.0 4 8 US-08-681-219-20
41 4 100.0 4 8 US-08-484-409-4
42 4 100.0 4 8 US-08-484-409-5
43 4 100.0 4 8 US-08-484-409-6
44 4 100.0 4 8 US-08-484-409-19
45 4 100.0 4 8 US-08-484-409-28

ALIGNMENTS

RESULT 1
US-09-982-172-145
; Sequence 145, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBIO
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE
; FILE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 145
; LENGTH: 2
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-145

Query Match 100.0% Score: 4; DB 10; Length 2;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
Db 1 V 1

RESULT 2
US-09-982-172-152
; Sequence 152, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBIO
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE
; FILE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172

;; CURRENT FILING DATE: 2001-10-19
;; NUMBER OF SEQ ID NOS: 253
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 152
;; LENGTH: 2
;; TYPE: PRT
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-152

Query Match 100.0%; Score 4; DB 10; Length 2;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 1 V 1

RESULT 3
US-08-424-550B-186
;; Sequence 186, Application US/08424550B
;; Patent No. US20020119447A1
;; GENERAL INFORMATION:
;; APPLICANT: JOHN N. SIMONS
;; APPLICANT: TAMI J. PILOT-MATIAS
;; APPLICANT: GEORGE J. DAWSON
;; APPLICANT: GEORGE G. SCHLAUDER
;; APPLICANT: SURESH M. DESAI
;; APPLICANT: THOMAS P. LEARY
;; APPLICANT: ANTHONY SCOTT MUEHROFF
;; APPLICANT: JAMES C. ERKER
;; APPLICANT: SHERI L. BUIJK
;; APPLICANT: ISA K. MUSHAWAR
;; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;; NUMBER OF SEQUENCES: 716
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;; STREET: 100 ABBOTT PARK ROAD
;; CITY: ABBOTT PARK
;; STATE: IL
;; COUNTRY: USA
;; ZIP: 60064-3500
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/424,550B
;; FILING DATE:
;; CLASSIFICATION: 435435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E.
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 5527.PC.01
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-938-2623
;; INFORMATION FOR SEQ ID NO: 186:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-424-550B-186

Query Match 100.0%; Score 4; DB 8; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 3 V 3

RESULT 4
US-09-826-290-464
;; Sequence 464, Application US/09826290
;; Patent No. US2002016468A1
;; GENERAL INFORMATION:
;; APPLICANT: Durham, L. Kathryn
;; APPLICANT: Friedman, David L.
;; APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri
;; APPLICANT: Kimmel, Lida H.
;; APPLICANT: Rajesh Bhikhu
;; APPLICANT: Potter, David M.
;; APPLICANT: Rohlff, Christian
;; APPLICANT: Silber, B. Michael
;; APPLICANT: Stiger, Thomas R.
;; APPLICANT: Sunderland, P. Trey
;; APPLICANT: Townsend, Robert Reid
;; APPLICANT: White, Frost
;; APPLICANT: Williams, Stephen A.
;; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
;; TITLE OF INVENTION: Uses Thereof, Including Diagnosis and Treatment of
;; TITLE OF INVENTION: Alzheimer's Disease
;; FILE REFERENCE: 2572-1-001 N2
;; CURRENT APPLICATION NUMBER: US/09/826,290
;; CURRENT FILING DATE: 2001-04-30
;; PRIOR APPLICATION NUMBER: US 60/194,504
;; PRIOR FILING DATE: 2000-04-03
;; PRIOR APPLICATION NUMBER: US 60/253,647
;; PRIOR FILING DATE: 2000-11-28
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 464
;; TYPE: PRT
;; ORGANISM: homo sapien
US-09-826-290-464

Query Match 100.0%; Score 4; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 3 V 3

RESULT 5
US-09-900-936-11
;; Sequence 11, Application US/09900936
;; Patent No. US20020165141A1
;; GENERAL INFORMATION:
;; APPLICANT: Rodgers, Kathleen
;; APPLICANT: dizelega, Gere
;; TITLE OF INVENTION: Methods for Promoting Dendritic Cell Proliferation
;; TITLE OF INVENTION: or Differentiation
;; FILE REFERENCE: 00-506-A
;; CURRENT APPLICATION NUMBER: US/09/900,936
;; CURRENT FILING DATE: 2001-07-09
;; NUMBER OF SEQ ID NOS: 50
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 11
;; LENGTH: 3
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence:AI1 (1-3)
US-09-900-936-11

Query Match 100.0%; Score 4; DB 9; Length 3;

Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 3 V 3

RESULT 6
US-09-828-272A-1
; Sequence 1, Application US/09828272A
; Publication No. US20020183255A1
; GENERAL INFORMATION:
; APPLICANT: LIPTON, James M
; APPLICANT: CATANIA, Anna P
; TITLE OF INVENTION: USE OF KPV TRIPEPTIDE FOR DERMATOLOGICAL DISORDERS
; FILE REFERENCE: 259/058
; CURRENT APPLICATION NUMBER: US/09/828,272A
; CURRENT FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Designed polypeptide with anti-inflammatory, anti-bacterial, anti-
; OTHER INFORMATION: -fungal and antipyrretic properties.
US-09-828-272A-1

Query Match 100.0%; Score 4; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 3 V 3

RESULT 7
US-10-074-956-5
; Sequence 5, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-5

Query Match 100.0%; Score 4; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 3 V 3

RESULT 8
US-10-039-876A-6
; Sequence 6, Application US/10039876A
; Publication No. US20030032792A1
; GENERAL INFORMATION:

; APPLICANT: Conklin, Darrell C.
; APPLICANT: Blumberg, Hal
; TITLE OF INVENTION: A HUMAN 2-19 PROTEIN HOMOLOGUE, Z219A
; FILE REFERENCE: 97-63C1
; CURRENT APPLICATION NUMBER: US/10/039,876A
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/061,712
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: US 09/167,513
; PRIOR FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Motif 4, corresponding to residues 188 to 190 of
; OTHER INFORMATION: SEQ ID NO:2
US-10-039-876A-6

Query Match 100.0%; Score 4; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 2 V 2

RESULT 9
US-10-015-979-63
; Sequence 63, Application US/10015979
; Publication No. US20030036050A1
; GENERAL INFORMATION:
; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Verhaert, Raymond M.D.
; APPLICANT: Beekwilder, Martinus J.
; APPLICANT: Ahle, Wolfgang
; TITLE OF INVENTION: Enzyme Selection
; FILE REFERENCE: 2183-5207US
; CURRENT APPLICATION NUMBER: US/10/015,979
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/NL00/00399
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/138,443
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CB4 domain of mutant 18
; NAME/KEY: SITE
; LOCATION: (1)..(3)
; OTHER INFORMATION:
US-10-015-979-63

Query Match 100.0%; Score 4; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 2 V 2

RESULT 10
US-10-015-979-71
; Sequence 71, Application US/10015979

; Publication No. US20030036050A1
; GENERAL INFORMATION:
; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Verhaert, Raymond M.D.
; APPLICANT: Beekwilder, Martinus J.
; APPLICANT: Aehle, Wolfgang
; TITLE OF INVENTION: Enzyme Selection
; FILE REFERENCE: 2183-5207US
; CURRENT APPLICATION NUMBER: US/10/015,979
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: PCT/NL00/00399
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/138,443
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 71
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; NAME/KEY: SITE
; LOCATION: (1)..(3)
; OTHER INFORMATION:
US-10-015-979-71

Query Match 100.0%; Score 4; DB 9; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
DB 3 V 3

RESULT 11
US-09-803-854-9
; Sequence 9, Application US/09803854
; Patent No. US20010046686A1
; GENERAL INFORMATION:
; APPLICANT: Albert J. Wong
; APPLICANT: David K. Moscatello
; APPLICANT: Allan Lipton
; APPLICANT: Kim E. Leitzel
; TITLE OF INVENTION: Sensitive Detection of Wild-Type and
; TITLE OF INVENTION: Mutant EGFR by Specific ELISA Assays in Any Biological
; FILE REFERENCE: W0901-NP003
; CURRENT APPLICATION NUMBER: US/09/803,854
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/188,424
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-803-854-9

Query Match 100.0%; Score 4; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
DB 3 V 3

RESULT 12
US-09-804-866-14

; Sequence 14, Application US/09804866
; Patent No. US20020001814A1
; GENERAL INFORMATION:
; APPLICANT: Chu, Ivan K.
; APPLICANT: Lau, Tai-Chu
; APPLICANT: Siu, K. W. Michael
; TITLE OF INVENTION: Sequencing of Peptides by Mass Spectrometry.
; FILE REFERENCE: 7933.208-US-01
; CURRENT APPLICATION NUMBER: US/09/804,866
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: US 60/193,208
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 14
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Bovine ubiquitin
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: X is Q or K
US-09-804-866-14

Query Match 100.0%; Score 4; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
DB 1 V 1

RESULT 13
US-09-771-192-11
; Sequence 11, Application US/09771192
; Patent No. US20020049162A1
; GENERAL INFORMATION:
; APPLICANT: Rodgers, Kathleen
; APPLICANT: diZerega, Gere
; TITLE OF INVENTION: Methods for Inhibiting Smooth Muscle Cell Proliferation
; FILE REFERENCE: 99-1043-A
; CURRENT APPLICATION NUMBER: US/09/771,192
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: AII (1-3)
US-09-771-192-11

Query Match 100.0%; Score 4; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
DB 3 V 3

RESULT 14
US-09-853-918-6
; Sequence 6, Application US/09853918
; Patent No. US20020068346A1
; GENERAL INFORMATION:
; APPLICANT: Krystek, Stanley R.
; APPLICANT: Sheriff, Steven
; APPLICANT: Witmer, Mark R.
; APPLICANT: Hollenbaugh, Diane L.
; APPLICANT: Yan, Ning

; APPLICANT: Mouravieff, Julie E.
; APPLICANT: Einspahr, Howard M.
; APPLICANT: Kish, Kevin
; TITLE OF INVENTION: MODIFIED INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: DB24NP
; CURRENT APPLICATION NUMBER: US/09/853,918
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,448
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-918-6

Query Match 100.0%; Score 4; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 2 V 2

RESULT 15
US-09-850-061A-31
; Sequence 31, Application US/09850061A
; Patent No. US20020094957A1
; GENERAL INFORMATION:
; APPLICANT: NORDSTEDT, Christer
; APPLICANT: NASLUND, Jan
; APPLICANT: THYBERG, Johan
; APPLICANT: TJERNBERG, Lars O.
; APPLICANT: TERNIUS, Lars
; TITLE OF INVENTION: PEPTIDE BINDING THE KLVEF-SEQUENCE OF AMYLOID-BETA
; FILE REFERENCE: 033315-002
; CURRENT APPLICATION NUMBER: US/09/850,061A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: US 09/095,106
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: PCT/SE96/01621
; PRIOR FILING DATE: 1996-12-09
; PRIOR APPLICATION NUMBER: SE 9504467-3
; PRIOR FILING DATE: 1995-12-12
; PRIOR APPLICATION NUMBER: US 60/009,386
; PRIOR FILING DATE: 1995-12-29
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Amyloidosis
US-09-850-061A-31

Query Match 100.0%; Score 4; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 1 V 1

Search completed: May 12, 2003, 15:59:10
Job time : 2.94 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:33:02 ; Search time 1.14 Seconds
(without alignments)
84.328 Million cell updates/sec

Title: US-09-909-348-4

Perfect score: 4

Sequence: 1 V 1

Scoring table: BLOSUM62

Gapop 10.0., Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	4	100.0	3	3	P00010		angiotensin-conver
2	4	100.0	4	2	A37832		phenol 2-monooxyge
3	4	100.0	4	2	A61300		22K superhelical D
4	4	100.0	4	2	B43848		cell surface adhes
5	4	100.0	4	2	T46627		hypothetical prote
6	4	100.0	4	2	S17255		ribosomal protein
7	4	100.0	4	2	A27897		glucan 1,4-alpha-g
8	4	100.0	4	2	T30569		hypothetical prote
9	4	100.0	4	2	I34357		schwannomin - mous
10	4	100.0	4	2	A40135		branched-chain-ami
11	4	100.0	5	2	A60521		glycogen phosphory
12	4	100.0	5	2	JN0862		peptidyl-dipeptida
13	4	100.0	5	2	C41225		copper resistance
14	4	100.0	5	2	B22565		R-phycocyanin al
15	4	100.0	5	2	P00009		angiotensin-conver
16	4	100.0	5	2	T14908		hypothetical prote
17	4	100.0	5	2	T14910		hypothetical prote
18	4	100.0	5	2	S55237		zinc-binding prote
19	4	100.0	5	2	A44692		fulvicin - giant Af
20	4	100.0	5	2	B61168		cocoonase (EC 3.4.
21	4	100.0	5	2	I50385		myosin light chain
22	4	100.0	5	2	S11127		phosphoprotein, bo
23	4	100.0	5	2	PT0278		ig heavy chain CRD
24	4	100.0	5	2	G44817		27.5 kda structura
25	4	100.0	5	2	I44817		27.5K structural p
26	4	100.0	5	2	E44817		27.5K structural p
27	4	100.0	5	2	C44817		28.5K structural p
28	4	100.0	5	2	A44817		28K structural pro
29	4	100.0	6	2	S66195		alcohol dehydrogen

30 4 100.0 6 2 S02617 alcohol dehydrogen
31 4 100.0 6 2 JN0861 peptidyl-dipeptida
32 4 100.0 6 2 A35890 RNA-directed DNA P
33 4 100.0 6 2 B34835 dnaA protein - Pse
34 4 100.0 6 2 A60986 N-formyl oligopept
35 4 100.0 6 2 S11024 hydrogensulfite re
36 4 100.0 6 2 P00008 angiotensin-conver
37 4 100.0 6 2 A61140 sperm acrosomal pr
38 4 100.0 6 2 A27696 contraction-inhibi
39 4 100.0 6 2 B27696 contraction-inhibi
40 4 100.0 6 2 H48394 glycoprotein compo
41 4 100.0 6 2 S78764 ribosomal protein
42 4 100.0 6 2 A19780 transferrin - bovi
43 4 100.0 6 2 PC4392 whey glycoprotein
44 4 100.0 6 2 A49792 acylaminoacyl-pept
45 4 100.0 6 2 B26206 alpha-1,4-glucan-p

ALIGNMENTS

RESULT 1

PQ0010

angiotensin-converting enzyme inhibitor (FLP-3) - common fig

N:Alternate names: ficus latex peptide 3

C:Species: Ficus carica (common fig)

C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C:Accession: PQ0010

R:Maruyama, S.; Miyoshi, S.; Tanaka, H.

A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.

A:Reference number: PQ0008

A:Accession: PQ0010

A:Molecule type: protein

A:Residues: 1-3 <MAR>

A:Experimental source: latex

C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 100.0%; Score 4; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1

Db 2 V 2

RESULT 2

A37832

phenol 2-monooxygenase (EC 1.14.13.7) chain P5 - Pseudomonas sp. (strain CF600) (frag

C:Species: Pseudomonas sp.

C:Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 23-Jun-1993

C:Accession: A37832

R:Powlowski, J.; Shingler, V.

J. Bacteriol. 172, 6834-6840, 1990

A:Title: In vitro analysis of polypeptide requirements of multicomponent phenol hydro

A:Reference number: A37832; MUID:91072231; PMID:2254259

A:Accession: A37832

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-4 <POW>

C:Keywords: oxidoreductase

Query Match 100.0%; Score 4; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1

Db 4 V 4

RESULT 3

A61300
 22K superhelical DNA-binding protein - Escherichia coli (fragment)
 C:Species: Escherichia coli
 C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
 C:Accession: A61300
 R:Kishi, F.; Ebina, Y.; Miki, T.; Nakazawa, T.; Nakazawa, A.
 J. Biochem. 92, 1059-1068, 1982
 A:Title: Purification and characterization of a protein from Escherichia coli which forms a DNA-protein complex
 A:Reference number: A61300; MUID:83082696; PMID:6294066
 A:Accession: A61300
 A:Molecule type: protein
 A:Residues: 1-4 <KIS>
 C:Comment: This protein resembles some of the histone-like protein of bacteria in amino acid sequence
 C:Keywords: DNA binding; monomer

Query Match 100.0%; Score 4; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
 |
 Db 4 V 4

RESULT 4
 B43848
 cell surface adhesin for heparan sulfate, 60K - Staphylococcus aureus (fragment)
 C:Species: Staphylococcus aureus
 C:Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 24-Feb-1995
 C:Accession: B43848
 R:Liang, O.D.; Ascencio, F.; Fransson, L.A.; Wadstrom, T.
 Infect. Immun. 60, 899-906, 1992
 A:Title: Binding of heparan sulfate to Staphylococcus aureus.
 A:Reference number: B43848; MUID:92176005; PMID:1541563
 A:Accession: B43848
 A:Molecule type: protein
 A:Residues: 1-4 <LIR>
 A:Note: sequence extracted from NCBI backbone (NCBIP:854444)

Query Match 100.0%; Score 4; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
 |
 Db 3 V 3

RESULT 5
 T46627
 hypothetical protein c4 - lobliolly pine
 C:Species: Pinus taeda (lobliolly pine)
 C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
 C:Accession: T46627
 R:Chang, S.; Purvey, J.; Funkhouser, E.A.; Newton, R.J.; Cairney, J.
 submitted to the EMBL Data Library, July 1995
 A:Description: Cloning of a chitinase homolog which lacks chitin binding sites and is deduced from cDNA
 A:Reference number: T46627
 A:Accession: T46627
 A:Molecule type: mRNA
 A:Residues: 1-4 <CH>
 A:Cross-references: EMBL:U31309; NID:g974285; PID:g974292
 A:Experimental source: strain s6PT2xs6PT3; 8 month seedlings

Query Match 100.0%; Score 4; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
 |
 Db 4 V 4

RESULT 6
 S17255
 ribosomal protein YmL1, mitochondrial, questionable - yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 A:Variety: strain 07173
 C:Date: 23-Apr-1993 #sequence_revision 14-Sep-1994 #text_change 09-May-1997
 C:Accession: S17255
 R:Grohmann, L.; Graack, H.R.; Kruff, V.; Choli, T.; Goldschmidt-Reisin, S.; Kitakawa, F.E.S. Lett. 284, 51-56, 1991
 A:Title: Extended N-terminal sequencing of proteins of the large ribosomal subunit from yeast
 A:Reference number: S17255; MUID:91285106; PMID:2060626
 A:Accession: S17255
 A:Molecule type: protein
 A:Residues: 1-4 <GRO>
 C:Comment: A coding region for this protein could not be identified in the genome of yeast
 C:Genetics:
 A:Genome: nuclear
 C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 100.0%; Score 4; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
 |
 Db 2 V 2

RESULT 7
 A27897
 glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - Aspergillus phoenicis (fragment)
 N:Alternate names: glucoamylase
 C:Species: Aspergillus phoenicis
 C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 06-Dec-1996
 C:Accession: A27897
 R:Inokuchi, N.; Takahashi, T.; Irie, M.
 J. Biochem. 90, 1055-1067, 1981
 A:Title: Purification and characterization of a minor glucoamylase from Aspergillus niger
 A:Reference number: A27897; MUID:82075730; PMID:6796572
 A:Note: Aspergillus saitoi
 A:Accession: A27897
 A:Molecule type: protein
 A:Residues: 1-4 <INO>
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 100.0%; Score 4; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
 |
 Db 2 V 2

RESULT 8
 T30569
 hypothetical protein - Emericella nidulans
 C:Species: Emericella nidulans, Aspergillus nidulans
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C:Accession: T30569
 R:Morrice, J.; Mackenzie, D.A.; Parr, A.J.; Archer, D.B.
 Curr. Genet. 34, 379-385, 1998
 A:Title: Isolation and characterisation of the acetyl-CoA carboxylase gene from Aspergillus nidulans
 A:Reference number: T30569; MUID:99087906; PMID:9871120
 A:Accession: T30569
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-4 <MOR>
 A:Cross-references: EMBL:Y15996; NID:e1285512; PID:e1218041; PIDN:CAA75927.1

Query Match 100.0%; Score 4; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 4 V 4

RESULT 9
I54357
Schwannomin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C:Accession: I54357
R:Huynh, D.P.; Nechiporuk, T.; Pulst, S.
Hum. Mol. Genet. 3, 1075-1079, 1994
A:Title: Alternative transcripts in the mouse neurofibromatosis type 2 (NF2) gene are co
A:Reference number: I54357; MUID:95072570; PMID:7981675
A:Accession: I54357
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4 <RES>
A:Cross-references: GB:L28838; NID:g454836; PIDN:AAA57150.1; PID:g601923
C:Genetics:
A:Gene: NF2

Query Match 100.0%; Score 4; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 1 V 1

RESULT 10
A40135
branched-chain-amino-acid transaminase (EC 2.6.1.42), mitochondrial - rat (fragment)
N:Alternate names: branched-chain-amino-acid aminotransferase, mitochondrial
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-May-1992 #sequence_revision 13-May-1992 #text_change 30-Sep-1993
C:Accession: A40135
R:Hutson, S.M.; Wallin, R.; Hall, T.R.
submitted to the Protein Sequence Database, March 1992
A:Reference number: A40135
A:Accession: A40135
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-4 <HUT>
C:Keywords: aminotransferase; mitochondrion

Query Match 100.0%; Score 4; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 1 V 1

RESULT 11
A60521
glycogen phosphorylase (EC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)
N:Alternate names: glycogen phosphorylase b
C:Species: Liza ramada
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 11-May-2000
C:Accession: A60521
R:Bonamusa, L.; Baanante, I.V.
Comp. Biochem. Physiol. B 95, 295-301, 1990
A:Title: Purification and characterization of glycogen phosphorylase B from skeletal mus
A:Reference number: A60521; MUID:90227907; PMID:2109669
A:Accession: A60521
A:Molecule type: protein

A:Residues: 1-5 <BON>
C:Superfamily: phosphorylase
C:Keywords: glycosyltransferase; hexosyltransferase; phosphorylase b kinase) #status expe
F:3/Binding site: phosphate (Ser) (covalent) (by phosphorylase b kinase)

Query Match 100.0%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 4 V 4

RESULT 12
JN0862
peptidyl-di-peptidase A inhibitory peptide C112 - striped bonito
C:Species: Sarda orientalis (striped bonito)
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
C:Accession: JN0862
R:Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.
Biosci. Biotechnol. Biochem. 57, 1743-1744, 1993
A:Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory
A:Reference number: JN0859; MUID:94080036; PMID:7764272
A:Accession: JN0862
A:Molecule type: protein
A:Residues: 1-5 <MAT>
A:Experimental source: intestine
C:Comment: The amino terminal tripeptide of this protein inhibits angiotensin I-conve
C:Superfamily: bradykinin-potentiating peptide
C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 100.0%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 4 V 4

RESULT 13
C41225
copper resistance protein - Pseudomonas syringae pv. tomato (fragment)
C:Species: Pseudomonas syringae pv. tomato
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 24-Jun-1993
C:Accession: C41225
R:Cha, J.S.; Cooksey, D.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 8915-8919, 1991
A:Title: Copper resistance in Pseudomonas syringae mediated by periplasmic and outer
A:Reference number: A41225; MUID:92020961; PMID:1924351
A:Accession: C41225
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <CHA>

Query Match 100.0%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 5 V 5

RESULT 14
B22565
R-phycoerythrin alpha-2 chain - red alga (Gastrocionium coulteri) (fragment)
C:Species: Gastrocionium coulteri
C:Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C:Accession: B22565
R:Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985

A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A:Reference number: A22565; MUID:85182601; PMID:3886644
A:Accession: B22565
A:Molecule type: protein
A:Residues: 1-5 <KLO>

Query Match 100.0%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 3 V 3

RESULT 15

PQ0009

angiotensin-converting enzyme inhibitor (PLP-2) - common fig
N:Alternate names: ficus latex peptide 2
C:Species: Ficus carica (common fig)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 08-Dec-1995
C:Accession: PQ0009
R:Maruyama, S.; Miyoshi, S.; Tanaka, H.
Agric. Biol. Chem. 53, 2763-2767, 1989
A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.

A:Reference number: PQ0008
A:Accession: PQ0009
A:Molecule type: protein
A:Residues: 1-5 <MAR>
A:Experimental source: latex
C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 100.0%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
|
Db 4 V 4

Search completed: May 12, 2003, 15:39:44
Job time : 2.14 secs

Result No.	Query			ID	Description
	Score	Match	Length DB		
1	4	100.0	4	1 EOSI_HUMAN	P03731 homo sapien
2	4	100.0	4	1 RM01_YEAST	P36515 saccharomyc
3	4	100.0	5	1 ET03_LITRU	P82099 litoria rub
4	4	100.0	5	1 ET04_LITRU	P82100 litoria rub
5	4	100.0	5	1 REIL_LITRU	P82070 litoria rub
6	4	100.0	5	1 RE32_LITRU	P82073 litoria rub
7	4	100.0	5	1 TPIS_CANFA	P54714 canis famil
8	4	100.0	5	1 UC22_MAZE	P80628 zea mays (m
9	4	100.0	6	1 ACPH_RABIT	P25154 oryctolagus
10	4	100.0	6	1 CIP1_MYTED	P13736 mytilus edu
11	4	100.0	6	1 CIP2_MYTED	P13737 mytilus edu
12	4	100.0	6	1 EI01_LITRU	P83096 litoria rub
13	4	100.0	7	1 CCFL_ENTFA	P20104 enterococcu
14	4	100.0	7	1 WWA1_ACHFU	P35919 achatina fu
15	4	100.0	7	1 WWA2_ACHFU	P35920 achatina fu
16	4	100.0	7	1 WWA3_ACHFU	P35921 achatina fu
17	4	100.0	8	1 ACT_CARMA	P80709 carcinus ma
18	4	100.0	8	1 AKHG_GRAYBI	P14086 gryllus bim
19	4	100.0	8	1 AXH_LIBAU	P25418 libellula a
20	4	100.0	8	1 ANG2_BOTJA	Q10582 bothrops ja
21	4	100.0	8	1 CAD1_ENTFA	P13268 enterococcu
22	4	100.0	8	1 CLP_THICU	P80488 thiobacillus
23	4	100.0	8	1 COW2_CONPU	P58785 conus purpu
24	4	100.0	8	1 CPD1_ENTFA	P13269 enterococcu
25	4	100.0	8	1 FUSS_FUSSO	P81010 fusarium so
26	4	100.0	8	1 HTFI_PERAM	Q04548 periplaneta
27	4	100.0	8	1 NS3_MYCTU	P81152 mycobacteri
28	4	100.0	8	1 PLP_BRANA	P81707 brassica na
29	4	100.0	8	1 PKK3_PERAM	P82618 periplaneta
30	4	100.0	8	1 UPAL_HUMAN	P30087 homo sapien
31	4	100.0	8	1 UFPA_HUMAN	P30096 homo sapien
32	4	100.0	8	1 VSLG_HSUVB	P81780 herpes simp
33	4	100.0	9	1 COXE_THYOB	P80975 thunnus obe

RX MEDLINE=91285106; PubMed=2060626;
 RA Grohmann L.; Graack H.-R.; Kruft V.; Choli T.; Goldschmidt-Reisin S.

RA Kitakawa M.;
 RT "Extended N-terminal sequencing of proteins of the large ribosomal
 subunit from yeast mitochondria";
 RL FEBS Lett. 284:51-56(1991).
 DR PIR; S17255; S17255.
 DR SGD; L0002681; MRPL1.
 KW Ribosomal protein; Mitochondrion.
 FT NON_TER 4 4
 SQ SEQUENCE 4 AA; 402 MW; 7771825D00000000 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
 Db 2 V 2

RESULT 3

E103_LITRU
 ID E103_LITRU STANDARD; PRT; 5 AA.
 AC P82099;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Electrin 3.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Skin secretion;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 Litoria electrica. Comparison with the skin peptides from Litoria
 rubella";
 RL Aust. J. Chem. 52:639-645(1999).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 KW Amphibian skin; Amidation.
 FT MOD_RES 5 5
 SQ SEQUENCE 5 AA; 630 MW; 668761F2C9A00000 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
 Db 2 V 2

RESULT 4

E104_LITRU
 ID E104_LITRU STANDARD; PRT; 5 AA.
 AC P82100;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Electrin 4.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Skin secretion;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 Litoria electrica. Comparison with the skin peptides from Litoria

RT rubella";
 RL Aust. J. Chem. 52:639-645(1999).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 KW Amphibian skin; Amidation.
 FT MOD_RES 5 5
 SQ SEQUENCE 5 AA; 616 MW; 61F2D1A059A00000 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
 Db 4 V 4

RESULT 5

RE11_LITRU
 ID RE11_LITRU STANDARD; PRT; 5 AA.
 AC P82070;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Rubellidin 1.1.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE AND MASS SPECTROMETRY.
 RC TISSUE-Skin secretion;
 RA Tyler M.J., Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog
 'Litoria rubella', the skin peptide profile as a probe for the study
 of evolutionary trends of amphibians";
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
 ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=598; METHOD=FAB.
 KW Amphibian skin.
 SQ SEQUENCE 5 AA; 598 MW; 6DD9C9CAB2A00000 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1
 Db 1 V 1

RESULT 6

RE32_LITRU
 ID RE32_LITRU STANDARD; PRT; 5 AA.
 AC P82073;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Rubellidin 3.2.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Skin secretion;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 Litoria electrica. Comparison with the skin peptides from Litoria

```

RT rubella."
RL Aust. J. Chem. 52:639-645(1999).
CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
CC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 570 MW; 71A9C9C862A00000 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 1 V 1

RESULT 7
TPIS_CANFA STANDARD; PRT; 5 AA.
AC P54714;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Fragment).
GN TP11.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RC TISSUE-Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate - glycerone
CC phosphate.
CC -1- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
DR HSC-2DPAGE; P54714; DOG.
DR InterPro: IPR000652; Triophos_ismrse.
DR PROSITE: PS00171; TIM; PARTIAL.
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW Pentose shunt.
FT NON_TER 1 1
FT NON_TER 5 5
SQ SEQUENCE 5 AA; 550 MW; 64444862C9A00000 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 2 V 2

RESULT 8
UC22_MAIZE STANDARD; PRT; 5 AA.
AC P80628;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 474)
DE (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade."

```

```

OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE-Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernellet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.1, ITS MW IS: 30.4 kDa.
DR Maize-2DPAGE; P80628; COLEOPTILE.
DR MaizeDB; 123954; -.
FT NON_TER 1 1
FT NON_TER 5 5
SQ SEQUENCE 5 AA; 654 MW; 72CB19C9C0300000 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 5 V 5

RESULT 9
ACPH_RABIT STANDARD; PRT; 6 AA.
ID P25154;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acylamino-acid-releasing enzyme (EC 3.4.19.1) (Acyl-peptide hydrolase)
DE (APH) (Acylaminoacyl-peptidase) (Fragment).
GN APH.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RC TISSUE-Muscle;
RX MEDLINE=9222120; PubMed=1807161;
RA Krishna R.G., Chin C.C.Q., Wolf F.;
RT "N-terminal sequence analysis of N alpha-acetylated proteins after
RT unblocking with N-acylaminoacyl-peptide hydrolase.";
RL Anal. Biochem. 199:45-50(1991).
CC -1- FUNCTION: THIS ENZYME CATALYZES THE HYDROLYSIS OF THE AMINO-
CC TERMINAL PEPTIDE BOND OF AN N-ACETYLATED PEPTIDE TO GENERATE
CC AN N-ACETYLATED AA AND A PEPTIDE WITH A FREE AMINO-TERMINUS.
CC IT PREFERENTIALLY CLEAVES OFF AC-ALA, AC-MET AND AC-SER.
CC -1- CATALYTIC ACTIVITY: Acylaminoacyl-peptide + H(2)O - acylamino acid
CC + peptide.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9C.
DR MEROPS; S09.004; -.
DR InterPro: IPR002471; Prol_endopep_ser.
DR PROSITE: PS00708; PRO_ENDOPEP_SER; PARTIAL.
KW Hydrolase; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT NON_TER 6 6
SQ SEQUENCE 6 AA; 775 MW; 6732D6C40B16F000 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 5 V 5

```

```

RESULT 10
CIP1_MYTED STANDARD; PRT; 6 AA.
ID CIP1_MYTED AC P13736;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Contraction-inhibiting peptide I (MIP I).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RC TISSUE-Pedal ganglion;
RX MEDLINE=88240357; PubMed=3377776;
RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
RT "Structures and actions of Mytilus inhibitory peptides.";
RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
CC -I- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN
MUSCLES.
CC -I- SIMILARITY: TO MIP II.
DR PIR; A27696;
KW Hormone; Amidation.
FT MOD_RES 6
SQ SEQUENCE 6 AA; 637 MW; 72C9C68775B81000 CRC64;
Query Match 100.0%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 V 1
DB 6 V 6

RESULT 11
CIP2_MYTED STANDARD; PRT; 6 AA.
ID CIP2_MYTED AC P13737;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Contraction-inhibiting peptide II (MIP II).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RC TISSUE-Pedal ganglion;
RX MEDLINE=88240357; PubMed=3377776;
RA Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
RT "Structures and actions of Mytilus inhibitory peptides.";
RL Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
CC -I- FUNCTION: INHIBITORY ACTION ON CONTRACTIONS IN SEVERAL MOLLUSCAN
MUSCLES.
CC -I- SIMILARITY: TO MIP I.
DR PIR; B27696;
KW Hormone; Amidation.
FT MOD_RES 6
SQ SEQUENCE 6 AA; 621 MW; 72C9C6876DD81000 CRC64;
Query Match 100.0%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 V 1
DB 6 V 6

RESULT 12
EI01_LITRU STANDARD; PRT; 6 AA.
ID EI01_LITRU AC P82096;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electrin 1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonolidae; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE-Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
Litoria electra. Comparison with the skin peptides from Litoria
rubella.";
RL Aust. J. Chem. 52:639-645(1999).
CC -I- SUBCELLULAR LOCATION: SECRETED.
KW Amphibian skin; Amidation.
FT MOD_RES 6
SQ SEQUENCE 6 AA; 792 MW; 5683704772C9A000 CRC64;
Query Match 100.0%; Score 4; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 V 1
DB 2 V 2

RESULT 13
CCF1_ENTFA STANDARD; PRT; 7 AA.
ID CCF1_ENTFA AC P20104;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CCF10.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=89008313; PubMed=3139658;
RA Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA Adsit J.C., Dunny G.M., Suzuki A.;
RT "Structure of cCF10, a peptide sex pheromone which induces
RT conjugative transfer of the Streptococcus faecalis tetracycline
RT resistance plasmid, PCF10.";
RL J. Biol. Chem. 263:14574-14578(1988).
CC -I- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC PIR; A30812; A30812.
KW Pheromone.
RN [1]
RP SEQUENCE 7 AA; 790 MW; 72C9D2C731B2C740 CRC64;
Query Match 100.0%; Score 4; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 V 1
DB 2 V 2

RESULT 14
WWA1_ACHFU STANDARD; PRT; 7 AA.
ID WWA1_ACHFU

```

AC P35919;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Wwamide-1.
 OS Achatina fulica (Giant African snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 OC Achatinacea; Achatinidae; Achatina.
 OX NCBI_TaxID=6530;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Ganglion;
 RX MEDLINE=93265912; PubMed=8495720;
 RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
 RT "Wwamide-1, -2 and -3: novel neuromodulatory peptides isolated from
 ganglia of the African giant snail, Achatina fulica.";
 RL FEBS Lett. 323:104-108(1993).
 CC -1- FUNCTION: EXHIBITS MODULATORY EFFECTS ON THE PERIPHERAL NERVOUS
 CC SYSTEM. INHIBITS ACTIVITY ON A CENTRAL NEURON.
 CC PIR: S33245; S33245.
 DR Neuropeptide; Amidation.
 KW MOD_RES 7
 FT MOD_RES 7
 SQ SEQUENCE 7 AA; 993 MW; 7362D5B69B041310 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
 Db 6 V 6

RESULT 15
 WWA2_ACHFU
 ID WWA2_ACHFU STANDARD; PRT; 7 AA.
 AC P35920;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Wwamide-2.
 OS Achatina fulica (Giant African snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 OC Achatinacea; Achatinidae; Achatina.
 OX NCBI_TaxID=6530;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Ganglion;
 RX MEDLINE=93265912; PubMed=8495720;
 RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
 RT "Wwamide-1, -2 and -3: novel neuromodulatory peptides isolated from
 ganglia of the African giant snail, Achatina fulica.";
 RL FEBS Lett. 323:104-108(1993).
 KW Neuropeptide; Amidation.
 FT MOD_RES 7
 SQ SEQUENCE 7 AA; 964 MW; 7362D5B686D32310 CRC64;

Query Match 100.0%; Score 4; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
 Db 6 V 6

Search completed: May 12, 2003, 15:36:41
 Job time : 1.58 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 15:32:37 ; Search time 2.26 Seconds
(without alignments)
91.171 Million cell updates/sec

Title: US-09-909-348-4

Perfect score: 4

Sequence: 1 V 1

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	100.0	4	11 Q08433	Q08433 rattus nor
2	4	100.0	7	2 Q07354	Q07354 synechococ
3	4	100.0	7	2 Q47029	Q47029 enterobacte
4	4	100.0	7	2 Q05056	Q05056 actinobacil
5	4	100.0	7	2 Q34028	Q34028 spingomona
6	4	100.0	7	2 P70804	P70804 azotobacter
7	4	100.0	7	2 Q54248	Q54248 streptomyce
8	4	100.0	7	2 P72081	P72081 nocardia la
9	4	100.0	7	8 Q95945	Q95945 saccharomyc
10	4	100.0	7	10 Q9C5B3	Q9C5B3 arabidopsis
11	4	100.0	7	10 P93233	P93233 lycopersico
12	4	100.0	7	10 P82445	P82445 nicotiana t
13	4	100.0	7	12 Q66205	Q66205 transmissib
14	4	100.0	7	12 Q9YVE3	Q9YVE3 human adeno
15	4	100.0	7	12 Q9YIR0	Q9YIR0 human adeno
16	4	100.0	7	12 Q9YIQ9	Q9YIQ9 human adeno

17	4	100.0	7	13 Q42564	Q42564 fugu rubrip
18	4	100.0	8	2 Q68485	Q68485 klebsiella
19	4	100.0	8	2 Q9S443	Q9S443 pseudomonas
20	4	100.0	8	2 Q9Z629	Q9Z629 buchnera ap
21	4	100.0	8	2 Q45615	Q45615 bacillus su
22	4	100.0	8	2 Q49534	Q49534 mycoplasma
23	4	100.0	8	2 Q44463	Q44463 rhizobacea
24	4	100.0	8	2 Q56429	Q56429 thermus the
25	4	100.0	8	2 Q56140	Q56140 streptococc
26	4	100.0	8	2 P72279	P72279 rhodococcu
27	4	100.0	8	2 Q56759	Q56759 xanthobacte
28	4	100.0	8	2 Q45889	Q45889 clostridium
29	4	100.0	8	2 P83158	P83158 anabaena sp
30	4	100.0	8	2 Q93SP2	Q93SP2 pseudomonas
31	4	100.0	8	2 Q934S4	Q934S4 thiobacilli
32	4	100.0	8	2 P83152	P83152 anabaena sp
33	4	100.0	8	2 Q56246	Q56246 thermophilli
34	4	100.0	8	2 Q9R4M3	Q9R4M3 enterococcu
35	4	100.0	8	2 Q9R3X0	Q9R3X0 planktothri
36	4	100.0	8	2 Q53790	Q53790 streptococc
37	4	100.0	8	3 Q9HDS4	Q9HDS4 aspergillus
38	4	100.0	8	3 Q05403	Q05403 saccharomyc
39	4	100.0	8	3 Q9URB9	Q9URB9 saccharomyc
40	4	100.0	8	4 Q9UHK1	Q9UHK1 homo sapien
41	4	100.0	8	4 Q9Y4X6	Q9Y4X6 homo sapien
42	4	100.0	8	4 Q9H4D3	Q9H4D3 homo sapien
43	4	100.0	8	4 Q15894	Q15894 homo sapien
44	4	100.0	8	4 Q15895	Q15895 homo sapien
45	4	100.0	8	4 Q15902	Q15902 homo sapien

ALIGNMENTS

RESULT 1

Q08433 PRELIMINARY; PRT; 4 AA.

ID Q08433
AC Q08433;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-JAN-1999 (TREMREL. 09, Last annotation update)
DE UDP-glucuronosyltransferase, microsomal (EC 2.4.1.17) (UDPGT)
DE (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GUNN;
RX MEDLINE=91282758; PubMed=1840486;
RA Sato H., Aono S., Kashiwamata S., Koitai O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT hyperbilirubinemic Gunn rat.";
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
CC -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC ENDOGENOUS COMPOUNDS.
CC -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC BETA-D-GLUCURONOSIDE.
CC -!- SUBCELLULAR LOCATION: MICROSOME.
DR EMBL; S38636; AAB19259.1; -;
KW Transferase; Glycosyltransferase; Microsome; Multigene family.
FT NON_TER 1 1
FT NON_TER 4 4
SQ SEQUENCE. 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 100.0%; Score 4; DB 11; Length 4;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 V 1

```

Db      2 V 2
RESULT 2
O07354 PRELIMINARY; PRT; 7 AA.
AC O07354;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NifK (Fragment).
GN NIFK.
OS Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).
OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
OX NCBI_TaxID=14131;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RF-1;
RX MEDLINE=99231861; PubMed=10217509;
RT Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT "Organization and expression of nitrogen-fixation genes in the aerobic
RT nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT RF-1."
RL Microbiology 145:743-753(1999).
DR EMBL; AF003700; AAC35193.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 849 MW; 7412C72AA9D5B030 CRC64;

Query Match 100.0%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 6 V 6

RESULT 3
Q47029 PRELIMINARY; PRT; 7 AA.
AC Q47029;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE Aad A1 protein (Fragment).
GN Aad A1.
OS Enterobacter cloacae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
OX NCBI_TaxID=550;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94079349; PubMed=8257126;
RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
RT "Analysis of the aac(3)-Via gene encoding a novel 3-N-
RT acetyltransferase."
RL Antimicrob. Agents Chemother. 37:2074-2079(1993).
DR EMBL; M88012; AAA16193.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 100.0%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 4 V 4

RESULT 4
O50556 PRELIMINARY; PRT; 7 AA.
ID O50556

```

```

AC O50556;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GlyA (Fragment).
GN GLYA.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OS actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33384;
RX MEDLINE=96355846; PubMed=8751884;
RA Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,
RA Kraig E.;
RT "Cis Elements and trans factors are both important in strain-specific
RT regulation of the leukotoxin gene in Actinobacillus
RT actinomycetemcomitans."
RL Infect. Immun. 64:3451-3460(1996).
DR EMBL; U51862; AAB88721.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 100.0%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 5 V 5

RESULT 5
O34028 PRELIMINARY; PRT; 7 AA.
ID O34028;
AC O34028;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Catechol-2,3-dioxygenase (Fragment).
GN PHE.
OS Shingomonas chungbukensis.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC Shingomonas.
OX NCBI_TaxID=56193;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DJ77;
RA Kim Y.-C.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88298; AAB66311.1; -.
KW Dioxxygenase.
FT NON_TER 1
SQ SEQUENCE 7 AA; 868 MW; 71A452D1A699D460 CRC64;

Query Match 100.0%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
Db 5 V 5

RESULT 6
P70804 PRELIMINARY; PRT; 7 AA.
ID P70804
AC P70804;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

```

DE Algt protein (Fragment).
GN ALGT.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RX MEDLINE=96427318; PubMed=8930582;
RA Rehm B.H.A., Ertesvag H., Valla S.;
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is
RT part of an alg gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL: X87973; CAA61230.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 684 MW; 7185A5A5A2D1AED0 CRC64;
Query Match 100.0%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 V 1
Db 3 V 3
RESULT 7
Q54248 PRELIMINARY; PRT; 7 AA.
AC Q54248;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RplO protein (Fragment).
OS Streptomyces griseus.
GN Streptomyces griseus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2-3-11;
RX MEDLINE=20011291; PubMed=10542330;
RA Poshling S., Pipersberg W., Wehmeier U.F.;
RT "Analysis and regulation of the sec Y gene from Streptomyces griseus
RT N2-3-11 and interaction of the SecY protein with the SecA protein.";
RL Biochim. Biophys. Acta 1447:298-302(1999).
DR EMBL: X95915; CAA65160.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;
Query Match 100.0%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 V 1
Db 1 V 1
RESULT 8
P72081 PRELIMINARY; PRT; 7 AA.
AC P72081
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 3'-methylcephem hydroxylase (Fragment).
GN CEF.
OS Nocardia lactamurans.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96009872; PubMed=7557411;
RA Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
RA Liras P.;
RT "Characterization of the cmcH genes of Nocardia lactamurans and
RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
RT O-carbamoyltransferase for cephamycin biosynthesis.";
RL Gene 162:21-27(1995).
DR EMBL: Z21682; CAA79797.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 746 MW; 75A1B2CDD1E681C0 CRC64;
Query Match 100.0%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 V 1
Db 5 V 5
RESULT 9
Q95945 PRELIMINARY; PRT; 7 AA.
AC Q95945;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inside intron 5 (Fragment).
OS Saccharomyces cerevisiae (Baker's yeast).
GN Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D273-10B;
RX MEDLINE=81069885; PubMed=6254986;
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
RT "Assembly of the mitochondrial membrane system: Structure and
RT nucleotide sequence of the gene coding for subunit 1 of yeast
RT cytochrome oxidase.";
RL J. Biol. Chem. 255:11927-11941(1980).
DR EMBL: V00694; CAA24066.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;
Query Match 100.0%; Score 4; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 V 1
Db 3 V 3
RESULT 10
Q9C5B3 PRELIMINARY; PRT; 7 AA.
AC Q9C5B3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 0.7 kba protein (Fragment).
GN DIDI 10A-2B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21171025; PubMed=11277426;
 RA Vercauteren I., Van Der Schueren E., Van Montagu M., Cheysen G.;
 "Arabidopsis thaliana genes expressed in the early compatible
 interaction with root-knot nematodes.";
 RT Mol. Plant Microbe Interact. 14:288-299(2001).
 DR EMBL; AJ286350; CAB71014.2; -;
 KW Hypothetical protein.
 FT NON_TER 1
 FT NON_TER 7
 SQ SEQUENCE 7 AA; 719 MW; 6732C7287EB325D0 CRC64;
 Query Match 100.0%; Score 4; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6.6e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 V 1
 DB 6 V 6

RESULT 11
 ID P93233 PRELIMINARY; PRT; 7 AA.
 AC P93233;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14)
 DE (Fragment).
 GN LE-ACS1B.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97351561; PubMed=9207843;
 RA Ostler J.H., Olson D.C., Shiu O.Y., Yang S.F.;
 "Differential induction of seven 1-aminocyclopropane-1-carboxylate
 synthase genes by elicitor in suspension cultures of tomato
 (Lycopersicon esculentum).";
 RT Plant Mol. Biol. 34:275-286(1997).
 DR EMBL; U75692; AAC49682.1; -;
 KW Lyase.
 FT NON_TER 1
 FT NON_TER 7
 SQ SEQUENCE 7 AA; 828 MW; 71B412C7377415D0 CRC64;
 Query Match 100.0%; Score 4; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6.6e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 V 1
 DB 5 V 5

RESULT 12
 ID P82445 PRELIMINARY; PRT; 7 AA.
 AC P82445;
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE 10 kDa cell wall protein (Fragment).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.

OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. PETIT HAVANA;
 RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
 RA Wojtaszek P., Bolwell G.P.;
 "Proteomic study of secondary cell wall proteins from transformed
 tobacco culture.";
 RT Planta 0:0-0(2000).
 CC -1- SUBCELLULAR LOCATION: CELL WALL.
 CC -1- TISSUE SPECIFICITY: XYLEM.
 KW Cell wall.
 FT NON_TER 7
 FT NON_TER 7
 SQ SEQUENCE 7 AA; 758 MW; 69D2C1E862D1B2A0 CRC64;
 Query Match 100.0%; Score 4; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6.6e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 V 1
 DB 1 V 1

RESULT 13
 ID Q66205 PRELIMINARY; PRT; 7 AA.
 AC Q66205;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Membrane protein (1 is 3rd base in codon) (Fragment).
 OS Transmissible gastroenteritis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=11149;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FS772/70;
 RA MEDLINE=88216185; PubMed=2835592;
 RA Britton P., Carmenes R.S., Page K.W., Garwes D.J., Parra F.;
 "Sequence of the Nucleoprotein Gene from a Virulent British Field
 Isolate of Transmissible Gastroenteritis Virus and its Expression in
 Saccharomyces Cerevisiae.";
 RT Mol. Microbiol. 2:89-99(1988).
 DR EMBL; Y00542; CAA6806.1; -;
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 869 MW; 72C691E727233B70 CRC64;
 Query Match 100.0%; Score 4; DB 12; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6.6e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 V 1
 DB 7 V 7

RESULT 14
 ID Q9YVE3 PRELIMINARY; PRT; 7 AA.
 AC Q9YVE3;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PVI core protein (Fragment).
 GN PVI.
 OS Human adenovirus type 7.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=10519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GOMEN;

RX MEDLINE-99175282; PubMed-10074533;
 RA Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;
 RT "Strain variation in adenovirus serotypes 4 and 7a causing acute
 respiratory disease."; J. Clin. Microbiol. 37:1107-1112(1999).
 RL J. Clin. Microbiol. 37:1107-1112(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GOMEN;
 RA Crawford-Miksza L.K.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF065085; AAD03662.1; -;
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 100.0%; Score 4; DB 12; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6.6e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
 |
 Db 1 V 1

RESULT 15

Q9YIRO PRELIMINARY; PRT; 7 AA.
 AC Q9YIRO;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE PVI core protein (Fragment).
 GN PVI.
 OS Human adenovirus type 7a.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID-85755;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KN T96-0620, S-1058, AND CL 68578;
 RA Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;
 RT "Molecular surveillance of strain variation in adenoviruses causing
 acute respiratory disease, AV 4 and AV 7a."; J.
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF065068; AAD03668.1; -;
 DR EMBL; AF065066; AAD03664.1; -;
 DR EMBL; AF065067; AAD03666.1; -;
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 100.0%; Score 4; DB 12; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6.6e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 V 1
 |
 Db 1 V 1

Search completed: May 12, 2003, 15:38:45
 Job time : 4.26 secs